THE ACCOUNTING REVIEW

VOLUME 78, No. 1 JANUARY 2003

18.07.03

FRANK HEFLIN, K. R. SUBRAMANYAM, and YUAN ZHANG
Regulation FD and the Financial Information Environment: Early Evidence

JON S. DAVIS, GARY HECHT, and JON D. PERKINS
Social Behaviors, Enforcement, and Tax Compliance Dynamics

SUNIL DUTTA
Capital Budgeting and Managerial Compensation: Incentive and Retention Effects

JOSEPH V. CARCELLO and TERRY L. NEAL
Audit Committee Characteristics and Auditor Dismissals following "New" Going-Concern Reports

ANNE BEATTY and JOSEPH WEBER
The Effects of Debt Contracting on Voluntary Accounting Method Changes

JOHN F. BOSCHEN, AUGUSTINE DURU, LAWRENCE A. GORDON, and KIMBERLY J. SMITH
Accounting and Stock Price Performance in Dynamic CEO Compensation Arrangements

DAVIT ADUT, WILLIAM H. CREADY, and THOMAS J. LOPEZ
Restructuring Charges and CEO Cash Compensation: A Reexamination

CRISTI A. GLEASON and CHARLES M. C. LEE
Analyst Forecast Revisions and Market Price Discovery

MICHAEL B. CLEMENT and SENYO Y. TSE
Do Investors Respond to Analysts' Forecast Revisions as If Forecast Accuracy Is All That Matters?

DAVID MANRY, SAMUEL L. TIRAS, and CLARK M. WHEATLEY
The Influence of Interim Auditor Reviews on the Association of Returns with Earning

MARLENE A. PLUMLEE
The Effect of Information Complexity on Analysts' Use of That Information

LESLIE HODDER, MARY LEA McANALLY, and CONNIE D. WEAVER
The Influence of Tax and Nontax Factors on Banks' Choice of Organizational Form

STEPHEN P. BAGINSKI and JAMES M. WAHLEN
Residual Income Risk, Intrinsic Values, and Share Prices

PATRICIA M. FAIRFIELD, J. SCOTT WHISENANT, and TERI LOMBARDI YOHN
Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing

---

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2002–2003 Executive Committee of the Association are as follows:

President	G. Peter Wilson, Boston College
President-Elect	William L. Felix, Jr., The University of Arizona
Vice President	William E. (Bill) McCarthy, Michigan State University
Vice President–Education	Kevin D. Stocks, Brigham Young University
Vice President–Finance	Judy D. Rayburn, University of Minnesota
Vice President–International	Judy S. L. Tsui, The Hong Kong Polytechnic University
Vice President–Professional Relations	Samuel A. Vitkoski, BDO Seidman
Vice President–Publications	Lawrence A. Tomassini, The Ohio State University
Vice President–Research	Zoe-Vonna Palmrose, University of Southern California
Vice President–Sections and Regions	Joanna L. Ho, University of California, Irvine
Past President	Joel S. Demski, University of Florida

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive *Accounting Education News*, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates: P17275

Full membership:
The membership dues are based on the selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.*
(This does not include Section journals.)

Select one Association-wide journal	U.S. $85.00
Select two Association-wide journals	U.S. $95.00
All three Association-wide journals	U.S. $105.00

Associate membership for students residing in the United States and Canada:

Select one Association-wide journal	U.S. $20.00
Select two Association-wide journals	U.S. $30.00
All three Association-wide journals	U.S. $40.00

Electronic journal option available for an additional U.S. $20.00 to above-mentioned full-membership and associate-membership dues.

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Artificial Intelligence/Emerging Technologies; Auditing; Financial Accounting and Reporting; Gender Issues in Accounting; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Teaching and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be sent to Professor Terry Shevlin, Editor, *The Accounting Review*, Department of Accounting, University of Washington, 231 Mackenzie Hall/SOBA, Box 353200, Seattle, WA 98195-3200.

The submission fee of $75.00 for members of the AAA or $100.00 for nonmembers must accompany manuscripts when submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

The Accounting Review (ISSN 0001-4826) is published four times a year in January, April, July, and October by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rate of U.S. $175; membership subscription rate is $25 with discounts available; and emeritus subscription rate is $8.75. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to ***The Accounting Review,*** 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2003 American Accounting Association. All rights reserved. Printed in the United States of America.

CUlc - H02582 - 30 - P017275

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

Editor
Linda Smith Bamber
Consulting Editor
E. Michael Bamber
Associate Editors
Greg Clinch
Raffi J. Indjejikian
Steven J. Kachelmeier
Kenneth J. Klassen
Marlys Gascho Lipe
Joan L. Luft
Mark W. Nelson
Grace Pownall
Richard C. Sansing
Daniel B. Thornton
Senyo Tse

30

VOL. 78 JANUARY 2003 No. 1

FRANK HEFLIN, K. R. SUBRAMANYAM, and YUAN ZHANG
Regulation FD and the Financial Information Environment: Early Evidence 1
JON S. DAVIS, GARY HECHT, and JON D. PERKINS
Social Behaviors, Enforcement, and Tax Compliance Dynamics 39
SUNIL DUTTA
Capital Budgeting and Managerial Compensation: Incentive and Retention Effects 71
JOSEPH V. CARCELLO and TERRY L. NEAL
Audit Committee Characteristics and Auditor Dismissals following "New" Going-Concern Reports 95
ANNE BEATTY and JOSEPH WEBER
The Effects of Debt Contracting on Voluntary Accounting Method Changes 119
JOHN F. BOSCHEN, AUGUSTINE DURU, LAWRENCE A. GORDON, and KIMBERLY J. SMITH
Accounting and Stock Price Performance in Dynamic CEO Compensation Arrangements 143
DAVIT ADUT, WILLIAM H. CREADY, and THOMAS J. LOPEZ
Restructuring Charges and CEO Cash Compensation: A Reexamination 169
CRISTI A. GLEASON and CHARLES M. C. LEE
Analyst Forecast Revisions and Market Price Discovery 193
MICHAEL B. CLEMENT and SENYO Y. TSE
Do Investors Respond to Analysts' Forecast Revisions as if Forecast Accuracy Is All That Matters? 227
DAVID MANRY, SAMUEL L. TIRAS, and CLARK M. WHEATLEY
The Influence of Interim Auditor Reviews on the Association of Returns with Earnings 251
MARLENE A. PLUMLEE
The Effect of Information Complexity on Analysts' Use of That Information 275
LESLIE HODDER, MARY LEA McANALLY, and CONNIE D. WEAVER
The Influence of Tax and Nontax Factors on Banks' Choice of Organizational Form 297
STEPHEN P. BAGINSKI and JAMES M. WAHLEN
Residual Income Risk, Intrinsic Values, and Share Prices 327
PATRICIA M. FAIRFIELD, J. SCOTT WHISENANT, and TERI LOMBARDI YOHN
Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing 353
Editorial Data 372
Editorial Policy and Style Information 373

Editorial Staff of *The Accounting Review*
2001–2002

EDITOR
Linda Smith Bamber, University of Georgia
CONSULTING EDITOR
E. Michael Bamber, University of Georgia
ASSOCIATE EDITORS
Greg Clinch, Melbourne Business School
Raffi J. Indjejikian, University of Michigan
Steven J. Kachelmeier, The University of Texas at Austin
Kenneth J. Klassen, University of Waterloo
Marlys Gascho Lipe, University of Oklahoma
Joan L. Luft, Michigan State University
Mark W. Nelson, Cornell University
Grace Pownall, Emory University
Richard C. Sansing, Dartmouth College
Daniel B. Thornton, Queen's University
Senyo Tse, The University of Texas at Austin
BOOK REVIEW EDITOR
Jane L. Reimers, Florida State University

American Accounting Association

INTERIM EXECUTIVE DIRECTOR
Tracey E. Sutherland
COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Kathy Casper
Helen Griffith
Carol Stanton
Judy Cothern
Jan Kovarik (subcontracted)
OTHER AAA STAFF
Debora Gardner
Roland LaTulip
Patricia A. Hoban
Deirdre Harris
Mary Cole
James Szelminski
Linda Watterworth
Myriam Chacon
Barbara Brady

Editorial Advisory and Review Board

Ashiq Ali, The University of Arizona
Shannon W. Anderson, Rice University
Benjamin C. Ayers, University of Georgia
Frances L. Ayres, University of Oklahoma
William R. Baber, The George Washington University
Stephen P. Baginski, University of Georgia
Ramamurthy Balakrishnan, The University of Iowa
Orie E. Barron, Pennsylvania State University
Jan Barton, Emory University
Anne L. Beatty, Pennsylvania State University
Jacob G. Birnberg, University of Pittsburgh
Robert J. Bloomfield, Cornell University
Robert M. Bowen, University of Washington
Lawrence D. Brown, Georgia State University
David C. Burgstahler, University of Washington
Brian J. Bushee, University of Pennsylvania
Robert M. Bushman, The University of North Carolina at Chapel Hill
Carolyn M. Callahan, University of Arkansas
Youngsoon Susan Cheon, Chung-Ang University
Kevin C. W. Chen, Hong Kong University of Science & Technology
Chee W. Chow, San Diego State University
C. Bryan Cloyd, University of Illinois at Urbana–Champaign
Paul A. Copley, University of Georgia
Patricia M. Dechow, University of Michigan
Mark L. DeFond, University of Southern California
Ilia D. Dichev, University of Michigan
Sunil Dutta, University of California, Berkeley
Pieter T. Elgers, University of Massachusetts
Merle M. Erickson, University of Chicago
John (Harry) Evans III, University of Pittsburgh
Gerald A. Feltham, University of British Columbia
Paul E. Fischer, Pennsylvania State University
Joseph G. Fisher, Indiana University
Jere R. Francis, University of Missouri–Columbia
O. Finley Graves, Kansas State University
Wayne R. Guay, University of Pennsylvania
David A. Guenther, University of Colorado at Boulder
Mahendra Gupta, Washington University
Susan F. Haka, Michigan State University
John R. M. Hand, The University of North Carolina at Chapel Hill
Donald R. Herrman, Oregon State University
D. Eric Hirst, The University of Texas at Austin
Chris E. Hogan, Southern Methodist University
Patrick E. Hopkins, Indiana University
K. E. (Skip) Hughes II, Louisiana State University
Christopher D. Ittner, University of Pennsylvania
Debra C. Jeter, Vanderbilt University
James J. Jiambalvo, University of Washington
Deen Kemsley, Columbia University
S. Jane Kennedy, University of Washington
William R. Kinney, Jr., The University of Texas at Austin
Lisa L. Koonce, The University of Texas at Austin
William N. Lanen, University of Michigan
Charles M. C. Lee, Cornell University
Robert Libby, Cornell University
Robert C. Lipe, University of Oklahoma
Russell J. Lundholm, University of Michigan

Joseph Magliolo III, Southern Methodist University
Laureen A. Maines, Indiana University
Ella Mae Matsumura, University of Wisconsin–Madison
Steven R. Matsunaga, University of Oregon
Linda S. McDaniel, University of Kentucky
Maureen F. McNichols, Stanford University
Richard R. Mendenhall, University of Notre Dame
Lillian F. Mills, The University of Arizona
James N. Myers, University of Illinois at Urbana–Champaign
Patricia C. O'Brien, University of Waterloo
Mark E. Peecher, University of Illinois at Urbana–Champaign
Kathy R. Petroni, Michigan State University
Jamie H. Pratt, Indiana University
Robert J. Ramsay, University of Kentucky
Madhav V. Rajan, Stanford University
Stephen G. Ryan, New York University
Katherine Schipper, Financial Accounting Standards Board
D. Shores, University of Washington
Paul J. Simko, University of Virginia
Richard G. Sloan, University of Michigan
J. Reed Smith, SUNY at Buffalo
Geoffrey B. Sprinkle, Indiana University
Thomas L. Stober, University of Notre Dame
Steve G. Sutton, University of Connecticut/University of Melbourne
Hun-Tong Tan, Nanyang Technological University
Siew Hong Teoh, The Ohio State University
Lynda Thoman, Purdue University
Richard Tubbs, The University of Iowa
Igor Vaysman, INSEAD
Robert E. Verrecchia, University of Pennsylvania
Martha L. Wartick, University of Northern Iowa
Gregory B. Waymire, Emory Univeristy
Hong Xie, University of Illinois at Urbana–Champaign
Teri L. Yohn, Georgetown University

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1960–1962	Robert K. Mautz
1963–1964	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William R. Kinney, Jr.
1990–1994	A. Rashad Abdel–khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 1–37

Regulation FD and the Financial Information Environment: Early Evidence

Frank Heflin
Purdue University
K. R. Subramanyam
Yuan Zhang
University of Southern California

ABSTRACT: On October 23, 2000, the SEC implemented Regulation FD (Fair Disclosure), which prohibits firms from privately disclosing value-relevant information to select securities markets professionals without simultaneously disclosing the same information to the public. We examine whether Regulation FD's prohibition of selective disclosure impairs the flow of financial information to the capital markets prior to earnings announcements. After implementation of FD, we find (1) improved informational efficiency of stock prices prior to earnings announcements, as evidenced by smaller deviations between pre- and post-announcement stock prices; (2) no reliable evidence of change in analysts' earnings forecast errors or dispersion; and (3) a substantial increase in the volume of firms' voluntary, forward-looking, earnings-related disclosures. Overall, we find no evidence Regulation FD impaired the information available to investors prior to earnings announcements, and some of our evidence is consistent with improvement.

Keywords: *disclosure; regulation; earnings announcements; capital markets; earnings forecasts; financial analysts.*

Data Availability: *All data employed in this study are commercially available from the sources described in the text.*

We are grateful to First Call/Thomson Financial for providing us with data used in this study. We thank the SEC and Financial Reporting Institute at the Leventhal School of Accounting, University of Southern California and the Krannert Graduate School of Management, Purdue University for their generous financial support. We also acknowledge Jinyoung Park's able research assistance. Finally, we thank two anonymous referees, Mark L. DeFond, William J. Kross, Jonathan Shokobin of the SEC, Frank Fernandez of the Securities Industry Association, and participants at the Financial Management Association Conference, Toronto, and research workshops at the University of Colorado at Boulder, University of Maryland, and Mellon Capital Management for their comments and suggestions. All remaining errors are our own.

Submitted July 2001
Accepted August 2002

I. INTRODUCTION

On October 23, 2000, the voluntary disclosure practices of corporations with publicly traded securities became subject to the requirements of the Securities and Exchange Commission's (SEC's) Regulation Fair Disclosure (FD). Regulation FD prohibits corporations from privately disclosing material information to select investors or securities markets professionals without simultaneously disclosing the same information to the public. Regulators intended FD to eliminate superior trading opportunities for recipients of firms' selective disclosures.[1] However, many securities markets professionals contend FD reduced the quantity and quality of information available to the capital markets, resulting in less accurate expectations of firm performance and larger price shocks when firms announce earnings (e.g., Securities Industry Association 2001; Association for Investment Management and Research 2001). In this paper, we empirically investigate whether the implementation of FD is associated, *on average*, with changes in the earnings-related information environment.[2] Specifically, we test for changes in (1) the informational efficiency of stock prices prior to firms' quarterly earnings announcements; (2) the accuracy and dispersion of analysts' earnings forecasts; and (3) the frequency of firms' voluntary disclosures. In general, we find no evidence the information available to investors prior to earnings announcements deteriorated after FD. On the contrary, we find some improvement in stock price efficiency and a substantial increase in the number of voluntary disclosures after FD's implementation.

The SEC implemented Regulation FD to address the commission's (particularly former chairman Arthur Levitt's) concern that firms communicate value-relevant information to select professional analysts and investors before disclosing it to the public.[3] Selective disclosure provides select investors an informational advantage from which they can profit at the expense of others, and the SEC was concerned this practice undermines investor confidence in the integrity of the capital markets. Regulation FD now requires that, when a firm's management intentionally discloses material information to select market participants, such as securities markets professionals or investors who may trade on the information, the firm must simultaneously make public that information. When a firm's management unintentionally discloses material information to select market participants, it must make that information public as soon as practical, but no later than 24 hours after the initial disclosure.

In contrast, analysts believe direct communication between management and analysts is a primary means by which corporations communicate performance-relevant information to the capital markets. The analysts' community suggests FD has impaired information available to investors, in part because FD prohibits firms from privately guiding analysts' earnings forecasts, which form the basis for investors' earnings expectations. For example, the director of research at Thomson Financial/First Call states, "You can't give guidance to individual analysts anymore; that will inevitably lead to wider ranges in estimates and more surprises" (Williams and McGough 2000). Additionally, FD may reduce the amount of detailed, performance-relevant information firms provide analysts, which also may impair their ability to forecast earnings and make buy-sell recommendations.[4] Furthermore, FD critics argue SEC 8-K filings and public announcements are insufficient substitutes for earnings guidance. When firms channel information through professional analysts, the analysts can guide the press, and ultimately investors, as to the information's appropriate context and meaning. However, since firms must now release information directly to the media, reporters and editors must draw inferences about context and meaning without the benefit of analysts' experience and expertise (Weber 2000). Additionally, FD critics suggest firms are less forthcoming in public announcements than in private conversations with analysts, partly because they fear litigation arising from improperly interpreted public announcements. Managers may also fear public disclosure of detailed information will benefit competitors, whereas professional analysts potentially use that information to improve earnings forecasts without disclosing competitively sensitive details (Opdyke and Lucchetti 2000; Weber 2000). Security Industry Association (SIA) spokesman Stuart Kaswell summarizes these concerns: "The playing field will be more level, but it will be empty" (Hassett 2000).

---

[1] See the U.S. Code of Federal Regulations 17 CFR Part 243.

[2] Although earnings-related information is not the only type of information FD addresses, it is clearly a primary target. For example, the regulation lists earnings information first among the types of information addressed and provides specific guidance on how to make a "planned disclosure of material information, such as an earnings release."

[3] A number of financial press articles, such as Pulliam (1999), report selective disclosure, and the text of FD refers to several similar reports. See the U.S. Code of Federal Regulations 17 CFR Part 243.

However, several factors may prevent deterioration in the information environment after the implementation of FD. First, analysts may be able to substitute information gathered from private search for information previously obtained directly from firms. For example, the majority of respondents to an Association for Investment Management and Research (AIMR) survey believe the accuracy of their earnings forecasts and the appropriateness of their stock recommendations will *not* suffer as a consequence of FD, despite reporting deterioration in private corporate communication to analysts (AIMR 2001). Second, companies may increase the quality and quantity of information dissemination through public disclosures. Finally, if private communication between firms and analysts cannot be (or is not) adequately monitored, FD may not successfully curtail selective disclosure to analysts. Therefore, the effect of FD on the financial information environment is an empirical question.

We analyze changes in various aspects of the financial information environment after the implementation of FD for a large sample of firms. Our test period consists of three post-FD quarters (the fourth quarter of 2000 and the first and second quarters of 2001). We compare each post-FD quarter to the latest like pre-FD quarter (the fourth quarter of 1999 and the first and second quarters of 2000, respectively). We first examine the effect of FD on the speed and extent to which stock prices anticipate information in upcoming earnings announcements. Specifically, we measure the "information gap" at various days prior to earnings announcements as the absolute deviation between the price on that day and the post-earnings-announcement stock price, after controlling for market-wide movements. A smaller information gap suggests the market has more information about the upcoming earnings announcement. The preponderance of our results suggests a *smaller* information gap over almost the entire quarter prior to earnings announcements, after FD. We next investigate the effect of FD on various aspects of analysts' forecasting performance. Although our univariate tests suggest that, on average, forecast accuracy declined and forecast dispersion increased after FD, regression analyses controlling for non-FD-related factors suggest FD had little effect on either forecast accuracy or forecast dispersion. Finally, we examine the effect of FD on the frequency of firms' voluntary public disclosures. We find a significant increase in voluntary earnings-related disclosures after FD. The increase in the number of firm-quarters with at least one disclosure is particularly significant.

---

[4] For example, 72 percent of analysts responding to a Securities Industry Association (SIA 2001) survey and 56 percent of analysts responding to an Association for Investment Management and Research (AIMR 2001) survey indicate that the "overall quality" of information companies disseminate has declined as a result of FD. A majority of analysts responding to the AIMR survey believe the accessibility and responsiveness of corporate officers, as well as general corporate communication with analysts, have declined since FD, and 47 and 57 percent of the SIA and AIMR survey respondents, respectively, believe the companies they follow now disseminate less total information.

In summary, we find no evidence that Regulation FD impaired the quality and quantity of investors' information prior to earnings announcements. We do not find deterioration in analysts' forecasting performance, but do find an improvement in the informational efficiency of prices prior to earnings announcements. We also find a marked increase in firms' voluntary disclosure frequency, which is consistent with firms substituting public disclosure for private communication through analysts.

Our results are potentially of interest to the SEC, which must assess the consequences of Regulation FD in a timely manner, especially given the criticism from the professional investment community. Our results are also potentially informative to the regulation's critics, largely securities markets professionals, since our systematic evidence suggests that their concerns that the regulation would have dire consequences for the financial information environment are likely unfounded. Finally, our results are potentially of interest to firms making disclosure decisions in the wake of this new regulation.

Our study provides early evidence on the cross-sectional average effects of FD on a broad range of issues related to the financial information environment. Recently, other researchers have explored the effect of FD on analysts' forecast performance in more detail. While broadly consistent with our results, this later work provides some interesting additional evidence. For example, Shane et al. (2001) find that errors in forecasts issued early in a quarter increase after FD, although there is no change for those issued later in the quarter.[5] Mohanram and Sunder (2002) find no change in analysts' forecast accuracy or dispersion, but find an increase in the idiosyncratic component of analysts' information (but no change in the common information component).

We organize the rest of the paper as follows. Section II describes the sample. Section III discusses changes in the informational efficiency of stock prices. Section IV presents analyses of analysts' forecasts. Section V presents analyses of firms' voluntary disclosures, and Section VI concludes.

II. SAMPLE AND DATA

We assess the effect of FD by analyzing, before and after FD, various proxies for firms' financial information environment prior to quarterly earnings announcements. To align calendar and fiscal quarters, we confine our sample to December fiscal year-end firms. Not all our sample firms announced 2000 third quarter earnings before FD's effective date (October 23, 2000). Hence, we exclude the third quarter of 2000 from our analysis and define the fourth quarter of 2000 and the first and second quarters of 2001 as our post-FD quarters. We match each post-FD quarter with its most recent analogous quarter prior to October 23, 2000. Accordingly, our pre-FD quarters are the fourth quarter of 1999 and the first and second quarters of 2000. By limiting our analyses to like quarters, we control for potential differences across quarters in stock price reactions to earnings (Mendenhall and Nichols 1988), analysts' forecast accuracy and dispersion, and voluntary disclosure frequency. Comparing quarters close in calendar time minimizes the risk that changes in other economic variables contaminate our results.

---

[5] Results in Shane et al. (2001) concerning errors from forecasts close to the earnings announcement are consistent with ours.

Our sample includes all December fiscal year-end firms with the following data available during our three pre- and three post-FD quarters: (1) current and previous quarters' earnings announcement dates; (2) current quarter and previous year's like-quarter actual earnings per share (EPS); (3) consensus EPS forecasts, where at least one analyst has updated his/her forecast since the previous quarter's earnings announcement; and (4) stock returns.[6] We obtain stock return data from CRSP and earnings announcement dates and actual and forecasted EPS data from First Call. Data for control variables, which we discuss in more detail later, are obtained from First Call, CRSP, or Compustat.

For each test, if a firm enters the sample in a post-FD quarter, we require that it also enter in the corresponding pre-FD quarter. Our final sample includes 5,072 pairs of pre- and post-FD observations from 2,025 distinct firms, with 1,669, 1,715, and 1,688 firms in the fourth, first, and second quarters, respectively. Missing data on one or more control variables causes us to use fewer observations in certain tests. This selection process biases our sample toward larger firms with relatively greater amounts of information available prior to their earnings announcements (Grant 1980; Atiase 1985; Freeman 1987; Kross and Schroeder 1989). Thus, our inferences may not apply to smaller companies with smaller amounts of publicly available information. Finally, to reduce the possibility our inferences are influenced by extreme observations, we winsorize all continuous variables at the 99th percentiles of the distributions of their absolute values.

III. EVIDENCE FROM THE INFORMATIONAL EFFICIENCY OF STOCK PRICES

In an efficient market, stock prices quickly and correctly reflect all value-relevant information. Accordingly, we begin our investigation of the effect of FD on information flows to the stock markets by analyzing the speed with which pre-announcement stock prices assimilate earnings information.

Earnings announcements reveal significant information about firm performance and move stock prices (Beaver 1968). However, the stock market anticipates much of this information and most of the price movement associated with earnings news *precedes* the earnings announcement (Ball and Brown 1968; Beaver et al. 1980). Superior pre-announcement information implies the market better anticipates earnings news and moves the price closer to its full information, post-announcement level (Freeman 1987).[7] Therefore, the absolute deviation between the price on any day prior to an earnings announcement and the post-announcement price is a measure of the information gap regarding the upcoming earnings announcement, with smaller deviations implying superior pre-announcement information.

---

[6] We eliminate all firms where First Call records "DDC," which denotes a discontinuity in the EPS series arising from events such as accounting changes, mergers, and acquisitions. Also, since it corresponds to the number in the earnings announcement, we use only the earliest First Call "actual" earnings number when the database contains, due to restatements, multiple earnings numbers.

[7] We assume the price two days after the announcement is the full information price because existing research (e.g., Morse 1981) shows prices impound most information by the day of (or after) an earnings announcement. Other research (e.g., Bernard and Thomas 1989, 1990; Freeman and Tse 1989) suggests the post-announcement price does not *fully* reflect the earnings announcement's information. We do not expect post-announcement drift to have any appreciable effect on our results because (1) the total earnings information-induced change in price occurring beyond two days after the announcement is a relatively small proportion of the total earnings-induced price change, and (2) evidence suggests the drift is due to investor naiveté regarding the magnitude of serial correlation in quarterly earnings (Ball and Bartov 1996) or researchers' difficulties specifying a quarterly earnings time-series model (Jacob et al. 2000). Neither likely has any relation to FD.

Accordingly, on each of the 64 trading days (the approximate number in a quarter) prior to the announcement we compute the absolute cumulative abnormal return as $ACAR_{i,q,x} = |\Pi_{t=-x}^{+2} (1 + AR_{i,q,t}) - 1|$ where x is the number of days the accumulation window extends *backward* from the earnings announcement date, and $AR_{i,q,t}$ is firm i's abnormal return on day t relative to the earnings announcement date of quarter q.[8] Abnormal returns are prediction errors from firm-specific estimations of the market model over the year ending the day before the start of the fiscal quarter.[9] We define the earnings announcement date as the day containing the announcement if the announcement is during trading hours, and the day after if it is after trading hours.[10]

$ACAR_{i,q,x}$ measures the absolute percentage change in price, after abstracting from market-wide movements, from x days before to two days after an earnings announcement. Thus, it measures the quarter's earnings news that is not reflected in stock price as of x days prior to the announcement (i.e., the information gap). Higher $ACAR_{i,q,x}$'s (ACARs) indicate larger information gaps. If, as critics suggest, FD reduced the flow of earnings-related information to market participants, *ceteris paribus,* we should observe higher ACARs after than before FD. Conversely, if FD reduced the information gap, ACARs should be lower for our post- than our pre-FD quarters.

Univariate Results

Figure 1 plots ACARs from 64 trading days before to two days after earnings announcements for both the pre- and post-FD quarters. Both the mean and median post-FD ACARs, in Panels A and B, respectively, are smaller than their pre-FD counterparts in the days leading up to and including the announcement day. The distance between the pre- and post-FD ACARs widens as time extends backward from the announcement, peaking at about day −30 and remaining relatively constant back to day −64. These plots suggest the information gap between the pre-announcement and the full-information post-announcement price is actually *smaller* post- than pre-FD.

The plots in Panels A and B display the *absolute* information flow for the pre- and post-FD quarters (i.e., they do not condition on the total information flow for the quarter). However, the post-FD plots lie below the pre-FD plots even at the beginning of the quarter, suggesting the *total* information flow over the quarter is smaller in our post- than pre-FD quarters. To measure the proportion of each quarter's total information impounded in price by any given day, we scale each day's mean or median ACAR by the mean or median, respectively, $ACAR_{i,q,-64}$, so that each plot starts the quarter at 1.0 and ends at 0.0. We

---

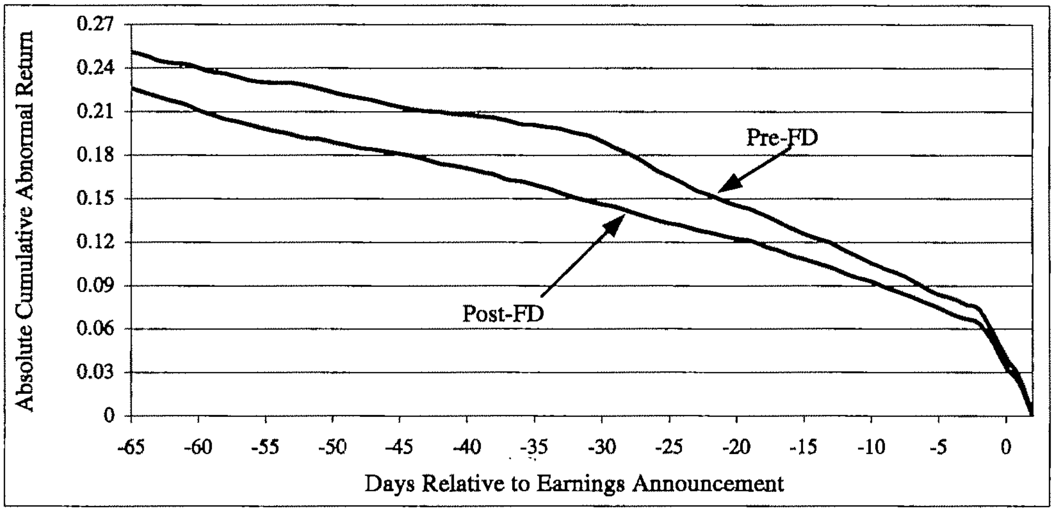
[8] Many studies (e.g., Beaver 1968; Landsman and Maydew 2002) use cumulative return variance (i.e., cumulative squared market model prediction errors) to examine stock price reactions to earnings announcements. We use $ACAR_{i,q,x}$ because it more directly measures the information "gap" and less likely produces extreme observations. While there is little difference between the two measures in short windows, they diverge over longer windows, since daily stock returns are negatively serially correlated (Lo and MacKinlay 1988). Other studies employing absolute returns include Morse (1981), Mendenhall and Fehrs (1999), Cheon et al. (2001), and Francis et al. (2002). Nevertheless, we replicate our four-, five-, and eight-day window analyses with squared market model prediction errors, both unscaled and scaled by the market model estimation period return variance. Our inferences are unchanged.

[9] Size-adjusted and market-adjusted returns yield qualitatively similar results.

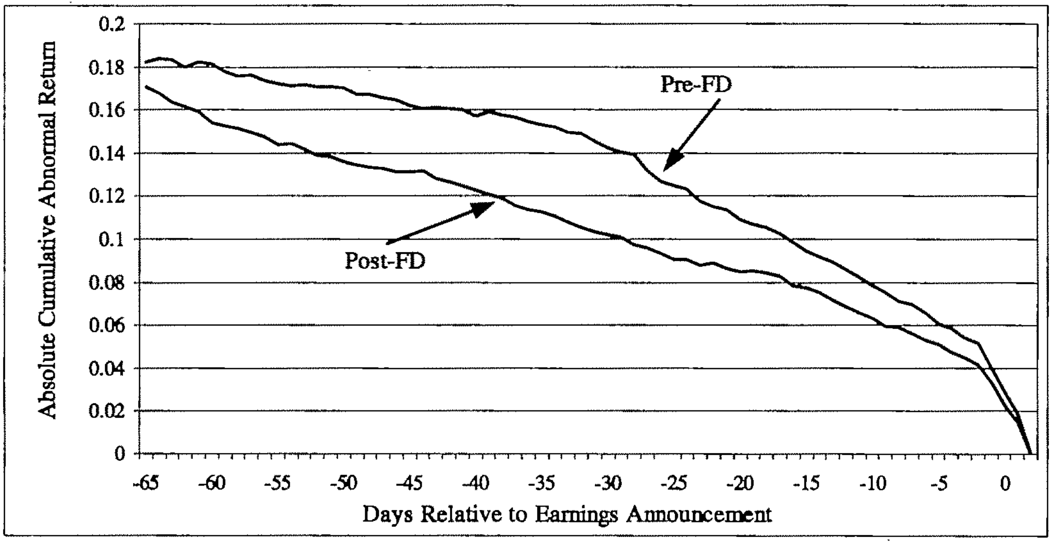
[10] We were able to match all but 137 of 2,472 First Call fourth-quarter 1999 earnings announcement dates to corresponding Compustat dates. For 127 of the 137 nonmatches, we located an earnings announcement on Dow Jones or one of the wire services available through Lexis/Nexis. The earliest wire service date coincided with the First Call date 90 times, the Compustat date 18 times, and with neither 19 times. This analysis mitigates concerns about the relative accuracy of the First Call announcement dates.

FIGURE 1
Price Discovery Before and After Regulation FD[a]

Panel A: Mean Absolute Cumulative Abnormal Return



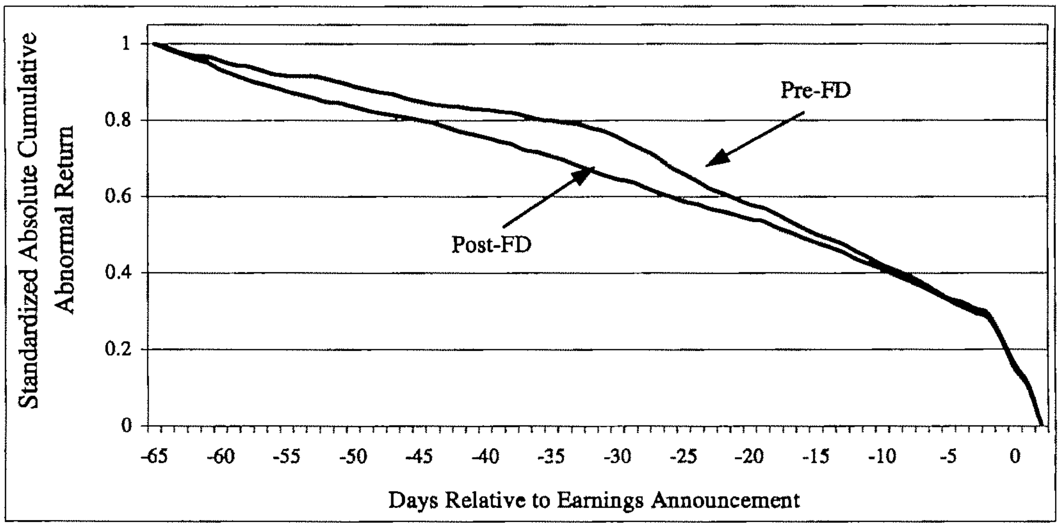
Panel B: Median Absolute Cumulative Abnormal Return



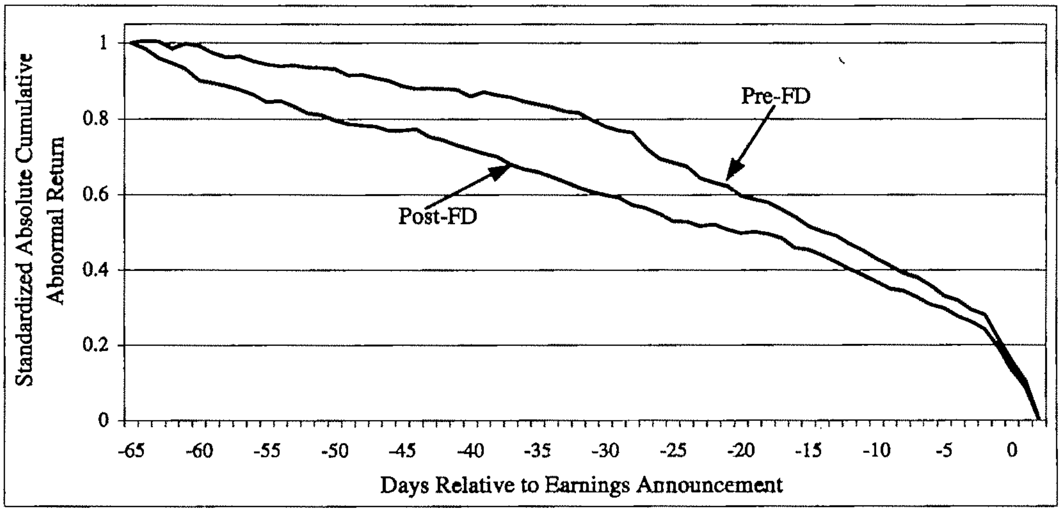
(*continued on next page*)

FIGURE 1 (continued)

Panel C: Daily Cross-Sectional Mean Absolute Cumulative Abnormal Return (ACAR) Scaled by Day –64 Cross-Sectional Mean ACAR



Panel D: Daily Cross-Sectional Median Absolute Cumulative Abnormal Return (ACAR) Scaled by Day –64 Cross-Sectional Median ACAR



---

[a] Each panel depicts either absolute, or standardized absolute, cumulative abnormal returns where the accumulation period begins on day −x (where −x is the value on the horizontal axis) and continues through day +2 relative to the earnings announcement date. Abnormal returns are prediction errors from the market model, estimated over the year ending the day before the start of the relevant fiscal quarter. Each variable is winsorized at the 99th percentile of the distribution of its absolute values.

present these standardized plots in Panels C and D of Figure 1.[11] In general, they are consistent with the plots in Panels A and B. For the entire 67-day period, both the mean and median standardized post-FD ACARs, in Panels C and D, lie below their pre-FD counterparts, suggesting smaller information gaps post-FD and general improvement in the information available to the market.

Table 1 presents statistical tests of differences between the pre- and post-FD pooled cross-sectional mean and median ACARs for accumulation windows of five different widths. The last day of each window is the second day after the earnings announcement (i.e., day +2) and the first day is alternatively day −1, −2, −5, −10, or −30 relative to the earnings announcement. The post-FD pooled cross-sectional mean and median ACARs are lower than their pre-FD counterparts for all five windows, and all differences are statistically significant.[12] The significant differences for the (−1,+2) and (−2,+2) windows suggest that, at the time of the quarterly earnings announcement, post-FD prices are closer to their full information levels than are pre-FD prices. The differences for the longer windows suggest that, even as many as 30 days prior to the announcement, post-FD prices are closer to their full information levels than pre-FD prices. Table 1 also presents the proportion of sample firms for which ACARs either decreased or increased from the pre- to post-FD periods. The percentage of firms with lower post-FD ACARs is significantly greater than the percentage with higher post-FD ACARs for every accumulation window. In summary, standard univariate comparisons of ACARs suggest that, post-FD, superior information is available to market participants prior to earnings announcements. Further, this superiority is pervasive through the entire quarter preceding the announcement.[13]

Simple comparisons of mean or median ACARs across a small number of quarters closely aligned in time can result in erroneous inferences because prior research reports large time-series variation in the price response to earnings announcements (Mendenhall and Fehrs 1999; Landsman and Maydew 2002). Accordingly, we also assess the significance of the differences in mean (and median) ACARs between pre- and post-FD quarters by benchmarking them to the historical time-series variation in such differences. Specifically, we gather a time series of 24 overlapping seven-consecutive-quarter sets during the period from the first quarter of 1993 through the second quarter of 2000. For each of the 24 sets of seven consecutive quarters, we label the first three quarters pseudo-pre-FD quarters and the last three quarters pseudo-post-FD quarters and exclude the middle quarter.[14] We then

---

[11] The standardized ACARs are conceptually appealing, but they suffer certain practical problems. First, small-denominator problems force us to scale each day's cross-sectional mean or median ACAR by the cross-sectional mean or median ACAR on day −64, rather than computing the mean or median of firm-specific scaled ACARs. Also, the scaled ACARs do not incorporate differences in informational efficiency *prior* to the current quarter.

[12] We report changes in cross-sectional means and medians, rather than mean and median within-firm changes to be consistent with our analysis in Section IV, where, as we explain later, the median within-firm change is not meaningful. However, analyses of mean and median within-firm changes from pre- to post-FD produce results very similar to the changes in means and medians reported in Table 1. Also, we do not report scaled average ACARs (as in Panels C and D of Figure 1) for brevity. However, scaled ACAR results are qualitatively similar to those reported in Table 1.

[13] Bamber et al. (2000) report that Beaver's (1968) U-statistic is greater than one for 25 to 30 percent of a sample of firms from the 1960s. We find U-statistics greater than 1 for the (−1,+2) window for 35 percent of our sample firms prior to FD and 23 percent after, consistent with the notion that information availability improved after FD.

[14] Our pseudo-pre- and pseudo-post-FD quarters are structured analogously to our actual pre- and post-FD quarters, i.e., three pre-FD quarters and three post-FD quarters separated by one excluded quarter in the middle. Also, as in our actual pre- and post-FD quarters, we require that any firm entering a pseudo-post-FD quarter must also enter its corresponding pseudo-pre-FD quarter, but we do not require that all firms have data for all pseudo-pre- and pseudo-post-FD quarters. Thus, while each pair of pseudo-pre- and corresponding pseudo-post-FD quarters employs the same firms, the set of firms used to compute the distributions for the Z-statistics changes over time.

TABLE 1
Absolute Cumulative Abnormal Returns ($ACAR_{i,q,x}$) and Pre- and Post-FD Earnings Announcements[a]

	−1,+2 Window ($ACAR_{i,q,-1}$)		*−2,+2 Window* ($ACAR_{i,q,-2}$)		*−5,+2 Window* ($ACAR_{i,q,-5}$)		*−10,+2 Window* ($ACAR_{i,q,-10}$)		*−30,+2 Window* ($ACAR_{i,q,-30}$)	
	Mean	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*
Pre-FD[b]	0.073	0.050	0.076	0.052	0.088	0.062	0.108	0.077	0.184	0.131
Post-FD	0.064	0.041	0.068	0.044	0.079	0.053	0.095	0.064	0.152	0.104
Difference:										
Magnitude	−0.008	−0.009	−0.009	−0.009	−0.009	−0.009	−0.013	−0.013	−0.032	−0.027
p-value[c]	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Z-statistic[d]	−1.690	−2.152	−1.532	−1.775	−1.293	−1.792	−1.514	−1.935	−2.410	−2.523
p-value	(0.09)	(0.03)	(0.13)	(0.08)	(0.20)	(0.07)	(0.13)	(0.05)	(0.02)	(0.01)
% of Firms with *Lower* Post-FD $ACAR_{i,q,x}$	55.26		54.24		53.86		55.03		56.70	
% of Firms with *Higher* Post-FD $ACAR_{i,q,x}$	44.74		45.76		46.14		44.97		43.30	
p-value	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	

---

[a] Variable definitions: $ACAR_{i,q,x}$ is the absolute cumulative abnormal return from x days before through 2 days after firm i's quarter q earnings announcement. Abnormal returns are prediction errors from the market model, estimated over the year ending the day before the start of the relevant fiscal quarter. We winsorize each $ACAR_{i,q,x}$ at the 99th percentile of the distribution of its absolute values.

[b] The fourth quarter of 1999 and the first and second quarters of 2000 are pre-FD quarters, while the fourth quarter of 2000 and the first and second quarters of 2001 are post-FD quarters.

[c] All p-values are two-sided. p-values for means are from t-tests of the difference between the pre- and post-FD means. For medians, p-values are from Wilcoxon two-sample tests; p-values for Z-statistics are from the student's t-distribution; p-values for differences in percentage of lower/higher post-FD ACARs are from binomial tests. There are 10,144 observations.

[d] We compute mean and median ACARs for each set of pseudo-pre- and pseudo-post-FD quarters in the 24 sets of seven consecutive quarters from the first quarter of 1993 through the second quarter of 2000. We label the first three quarters in each set pseudo-pre-FD quarters and the last three pseudo-post-FD quarters. The middle quarter is discarded. Z-statistics equal $(D_x - \overline{D}_x)/\sigma_x$, where D_x is the post-FD mean or median $ACAR_{i,q,x}$ minus the pre-FD mean or median $ACAR_{i,q,x}$ and $\overline{D}_x$ and σ_x are the mean and standard deviation, respectively, of the 24 differences between the pseudo-post- and pseudo-pre-FD means or medians.

compute Z-statistics equal to $(D_x - \overline{D}_x)/\sigma_x$ where D_x is the difference between the post- and pre-FD mean ACAR and $\overline{D}_x$ and σ_x are the mean and standard deviation, respectively, of the 24 pseudo-post- and pseudo-pre-FD differences in mean ACARs. Z-statistics for differences between pre- and post-FD medians are constructed analogously.

Table 1 reveals that, for the (−1,+2) window, the Z-statistic for the difference in mean ACARs is significant at the 0.09 level and for the difference in median ACARs is significant at the 0.03 level. Significance levels for the (−2,+2) window are marginal (p = 0.13 and 0.08 for the difference in mean and median, respectively). Z-statistics for the extended windows are mixed, with significance levels ranging from 0.20 for the (−5,+2) window difference in means to 0.01 for the (−30,+2) window difference in medians. Collectively, these univariate time-series tests reveal no evidence that Regulation FD impaired the flow of earnings-related information and, to the extent the pre- vs. post-FD differences are significant, continue to suggest an improvement in price efficiency, albeit weakly.[15]

Regression Results

In this section, we identify and control for factors other than Regulation FD that may cause differences in ACARs between our pre- and post-FD periods. In general, $ACAR_{i,q,x}$ is the product of the absolute unexpected earnings (absolute surprise) as of day −x and an earnings response coefficient (ERC). Holthausen and Verrecchia (1988) suggest three information-related factors that affect the ERC: inherent price variability, the precision of the market's information prior to the earnings announcement, and the precision of the earnings signal. In addition, empirical research has identified various non-information-related determinants of the ERC, such as interest rates and growth opportunities (e.g., Collins and Kothari 1989). Thus we could conceptually model ACAR as being determined by these four broad ERC determinants in addition to the (absolute) earnings surprise.

Our objective is to provide evidence on how FD affects the information available to the market prior to the earnings announcement, by observing changes in the ACARs. Accordingly, the treatment effect that we wish to capture includes any change in the ACARs due to earnings surprise attributable to FD and the effect of FD on the precision of the market's pre-announcement information. Thus, we need to control for any change in ACAR arising from earnings surprise unrelated to FD, inherent price variability, any aspect of the market's pre-announcement information precision unrelated to FD, the precision of the earnings signal, and other non-information-related determinants of the ERC. Accordingly, we identify control variables capturing these constructs. However, we do not control for the magnitude of unexpected earnings x days prior to the earnings announcement, because it is part of our treatment effect.

Our sample selection ensures each pre- and corresponding post-FD quarter contains the same firms, largely controlling for stationary firm-specific factors. We therefore focus primarily on controlling for factors affecting ACARs that likely changed from our pre- to post-FD periods but are unrelated to FD. Accordingly we estimate the following pooled cross-sectional and time-series regression:

---

[15] Recent research reports higher absolute abnormal returns at the time of earnings announcements in the 1990s than in previous decades (Francis et al. 2001, 2002; Kim and Kross 2002; Landsman and Maydew 2002). Results in Francis et al. (2002) suggest the increase is due to an increase in disclosures concurrent with earnings announcements. However, their Table 2 shows relatively constant abnormal returns around earnings announcements in the 1990s and it is not clear the trend they document extends to our very recent sample period. Our results are consistent with the notion that FD introduced a nonstationarity by changing the information environment.

$$\begin{aligned}\mathrm{ACAR}_{i,q,x} = {} & a_{0,x} + a_{1,x}\mathrm{POSTFD}_q + a_{2,x}\mathrm{RETVAR}_{i,q} \\ & + a_{3,x}\mathrm{NEGCAR}_{i,q} + a_{4,x}\mathrm{ABSCAR}_{i,q} \\ & + a_{5,x}\mathrm{LOSS}_{i,q} + a_{6,x}\mathrm{NEGSPEC}_{i,q} + a_{7,x}\mathrm{BOND30}_q \\ & + a_{8,x}\mathrm{EPRATIO}_{i,q} + e_{i,q,x}. \qquad (1)\end{aligned}$$

Our main variable of interest is POSTFD_q, which equals 1 if firm i's earnings announcement is from our post-FD quarters, and 0 otherwise. We expect the POSTFD_q coefficient to be positive if earnings-related information availability deteriorated after FD and negative if it improved.

We include $\mathrm{RETVAR}_{i,q}$ and $\mathrm{NEGCAR}_{i,q}$ as controls for inherent price variability. $\mathrm{RETVAR}_{i,q}$ is the standard deviation of firm i's abnormal returns during the market model estimation period for the relevant pre-FD quarter and controls for firm-specific inherent price variability. We expect it to be positively related to the ACARs (e.g., Beaver 1968).[16] $\mathrm{NEGCAR}_{i,q}$, which equals 1 if the cumulative abnormal return over the quarter is negative, and 0 otherwise, captures another dimension of price variability, as evidence suggests greater price movements in down markets than in up markets (Black 1976; Christie 1982).

We next include a set of variables designed to control for both the pre-announcement information precision and (absolute) earnings surprise unrelated to FD.[17] We include the absolute cumulative abnormal return during the entire quarter ($\mathrm{ABSCAR}_{i,q}$), as firm-quarters with larger total information flow are expected to have larger information gaps and larger ACARs at any given time.[18] We also include $\mathrm{LOSS}_{i,q}$ and $\mathrm{NEGSPEC}_{i,q}$ in this category of controls. $\mathrm{LOSS}_{i,q}$ equals 1 if firm i reports a loss in quarter q, and 0 otherwise, and $\mathrm{NEGSPEC}_{i,q}$ is the absolute value of special items deflated by total assets, if negative, and 0 otherwise.[19] We expect losses and large negative special items to impair the market's ability to forecast the upcoming earnings numbers. In addition to controlling for earnings surprise (or pre-announcement information precision), $\mathrm{LOSS}_{i,q}$ and $\mathrm{NEGSPEC}_{i,q}$ also capture differences in the precision of the earnings signal (Hayn 1995; Francis et al. 1996). While the relation of LOSS$_{i,q}$ and NEGSPEC$_{i,q}$ to multiple factors may make predicting and interpreting their coefficients more difficult, our interest is in the coefficient on POSTFD$_q$, not in the control variable coefficients.[20]

---

[16] Because $\mathrm{RETVAR}_{i,q}$ is constant for each pair of pre- and post-FD like quarters for each firm, its correlation with our variable of interest, POSTFD_c, is 0 by construction and, therefore, cannot affect POSTFD_q's estimated coefficient. However, we expect $\mathrm{RETVAR}_{i,q}$ to have considerable explanatory power for ACARs, so we include it to increase the power of our tests (Greene 1997, 403–404). Nonetheless, our inferences are unaffected if we exclude it. For brevity, we tabulate results from models excluding other (largely) time-invariant variables, although controlling for the natural log of equity market value, market model beta, and the book-to-price ratio does not change our inferences.

[17] Although (absolute) surprise and pre-announcement information precision are conceptually different constructs, it is difficult to empirically parse the two. Pre-announcement information precision is an *ex ante* construct that determines the magnitude of surprise on average, while surprise measures the *ex post* realization for a particular earnings announcement.

[18] Substituting the absolute value of the seasonal random-walk unexpected earnings for $\mathrm{ABSCAR}_{i,q}$ yields qualitatively similar results.

[19] We control for only negative special items because they are more common and of considerably greater magnitude than positive special items (e.g., Wild et al. 2001, 425). Furthermore, negative special items are frequently restructuring charges and asset impairments, which are both difficult to forecast. Finally, we find that the negative special-items variable we employ is more powerful (in our sample) than one indicating either positive or negative special items. Specifically, an indicator variable equaling 1 if the firm reports special items, irrespective of sign, is insignificant in all our models, in addition to having little effect on our POSTFD_q coefficients.

Our remaining control variables are designed to capture non-information-related determinants of the ERC. We include BOND30$_q$, the yield on the CRSP 30-year bond index at the end of the fiscal quarter, to capture changes in interest rates because prior research (e.g., Collins and Kothari 1989) suggests price reactions to earnings announcements are decreasing in interest rates. Finally, we include EPRATIO$_{i,q}$, which is the earnings-price ratio of firm i at the end of (fiscal) quarter q. We include EPRATIO$_{i,q}$ as a proxy for expected growth in earnings (Penman 1996), since prior research (e.g., Collins and Kothari 1989) argues that growth expectations increase stock price responses to earnings.

Panels A and B of Table 2 display descriptive information on the explanatory variables in Equation (1). The average standard deviation of daily returns (RETVAR$_{i,q}$) is 0.0345. Cumulative abnormal returns are negative for just over 51 percent of our firm-quarters (NEGCAR$_{i,q}$), reflecting the general decline in market values during the 1999–2001 period. The average absolute cumulative abnormal return (ABSCAR$_{i,q}$) is 24 percent and sample firms report losses in over 19 percent of our quarters (LOSS$_{i,q}$). The average 30-year bond yield during our sample period is 5.72 percent.

Panel A of Table 2 also shows means and medians for our control variables in both our pre- and post-FD periods. The post-FD means and medians are significantly different ($p = 0.01$ or better, two-sided) from their pre-FD counterparts for all control variables (except RETVAR$_{i,q}$, which, by construction, is the same for the three post- and three pre-FD quarters).[21] Thus, to the extent these variables are correlated with our variable of interest, POSTFD$_q$, their control is important. The correlation matrix, shown in Panel B of Table 2, reveals that POSTFD$_q$ is significantly correlated with all of our control variables, except RETVAR$_{i,q}$ ($p < 0.05$). However, with the exception of BOND30$_q$, the correlations are not economically large (less than 0.10 in magnitude). The correlation between POSTFD$_q$ and BOND30$_q$ is −0.7441, reflecting a decrease in long-term interest rates from the fourth quarter of 1999 through the second quarter of 2001. Also, most of the control variables are modestly but statistically significantly correlated with each other.

Panel C of Table 2 presents results from estimating Equation (1) for each of the five accumulation windows. The POSTFD$_q$ coefficient is negative and significant at better than the 0.01 level (two-sided) for all five accumulation windows, suggesting that the information gap actually declines after FD, as suggested by the univariate analyses in Table 1.

We also generate Z-statistics for the POSTFD$_q$ coefficients in an analogous manner to that for the differences in mean and median ACARs in our univariate analysis. Specifically, we estimate Equation (1) for each of the 24 sets of seven consecutive quarters from the first quarter of 1993 through the second quarter of 2000. POSTFD$_q$ equals 1 for the last three quarters in each seven-quarter set, and 0 for the first three quarters. We exclude the middle quarter. The Z-statistic equals $(\hat{a}_{1,x} - \bar{a}_{1,x})/\sigma_x$, where $\hat{a}_{1,x}$ is the estimate of $a_{1,x}$ (the

---

[20] Lower earnings precision (i.e., quality) unlikely explains the drop in price reactions to earnings announcements we report in Table 1. First, in our multiple regression tests, LOSS$_{i,q}$ and NEGSPEC$_{i,q}$ at least partially control for changes in earnings precision. Second, when we regress quarterly (long-window) returns on earnings, earnings change, and a loss indicator, we find higher R^2's after FD than before. These results indicate that earnings precision did not decline from our pre- to our post-FD period.

[21] In all our tables, p-values associated with means are from t-tests of the difference between the means, while p-values associated with medians are from Wilcoxon two-sample tests. The Wilcoxon two-sample test can indicate a significant difference between two samples even when the sample medians are identical (e.g., if the values above the median in sample A are generally larger than the values above the median in sample B, but the values below the medians of the two samples are roughly the same).

TABLE 2
Regression of Absolute Cumulative Abnormal Returns on Post-FD Indicator and Control Variables

$$ACAR_{i,q,x} = a_{0,x} + a_{1,x}POSTFD_q + a_{2,x}RETVAR_{i,q} + a_{3,x}NEGCAR_{i,q} + a_{4,x}ABSCAR_{i,q} + a_{5,x}LOSS_{i,q} + a_{6,x}NEGSPEC_{i,q} + a_{7,x}BOND30_q + a_{8,x}EPRATIO_{i,q} + e_{i,q,x}$$

Panel A: Distributional Statistics

Control	Means				Medians			
Variables[a]	*Overall*	*Pre-FD*	*Post-FD*	*p-value*[b]	*Overall*	*Pre-FD*	*Post-FD*	*p-value*
RETVAR	0.0345	0.0345	0.0345	NA	0.0304	0.0304	0.0304	NA
NEGCAR	0.5171	0.5430	0.4913	(0.00)	1.0000	1.0000	0.0000	(0.00)
ABSCAR	0.2441	0.2548	0.2334	(0.00)	0.1723	0.1785	0.1663	(0.01)
LOSS	0.1913	0.1706	0.2120	(0.00)	0.0000	0.0000	0.0000	(0.00)
NEGSPEC	0.0036	0.0022	0.0049	(0.00)	0.0000	0.0000	0.0000	(0.00)
BOND30	5.7226	5.9718	5.4734	(0.00)	5.7089	5.8048	5.3778	(0.00)
EPRATIO	0.0145	0.0158	0.0133	(0.00)	0.0128	0.0142	0.0118	(0.00)

Panel B: Pearson Correlations[c]

	RETVAR	*NEGCAR*	*ABSCAR*	*LOSS*	*NEGSPEC*	*BOND30*	*EPRATIO*
POSTFD	0.0000	−0.0518	−0.0376	0.0527	0.0854	−0.7441	−0.0954
RETVAR		0.1012	0.3084	0.4459	0.1264	0.0148	−0.3419
NEGCAR			−0.1253	0.0735	−0.0021	0.1115	−0.0852
ABSCAR				0.2146	0.0771	0.0729	−0.1421
LOSS					0.1487	−0.0273	−0.5354
NEGSPEC						−0.0252	−0.0995
BOND30							0.0410

(continued on next page)

TABLE 2 (continued)

Panel C: Regression Results

	(−1,+2)		(−2,+2)		(−5,+2)		(−10,+2)		(−30,+2)	
Variable[a]	*Coefficient*	*p-value*	*Coefficient*	*p-value*	*Coefficient*	*p-value*	*Coefficient*	*p-value*	*Coefficient*	*p-value*
Intercept	0.068	(0.00)	0.064	(0.00)	0.061	(0.00)	0.081	(0.00)	0.335	(0.00)
POSTFD	−0.012	(0.00)	−0.012	(0.00)	−0.011	(0.00)	−0.016	(0.00)	−0.053	(0.00)
RETVAR	1.207	(0.00)	1.236	(0.00)	1.304	(0.00)	1.449	(0.00)	1.475	(0.00)
NEGCAR	0.003	(0.04)	0.004	(0.02)	0.004	(0.02)	0.002	(0.31)	0.000	(0.99)
ABSCAR	0.040	(0.00)	0.051	(0.00)	0.078	(0.00)	0.120	(0.00)	0.336	(0.00)
LOSS	0.001	(0.70)	0.001	(0.65)	0.008	(0.00)	0.012	(0.00)	0.025	(0.00)
NEGSPEC	0.146	(0.00)	0.112	(0.02)	0.186	(0.00)	0.127	(0.05)	0.476	(0.00)
BOND30	−0.008	(0.01)	−0.007	(0.02)	−0.007	(0.06)	−0.009	(0.03)	−0.049	(0.00)
EPRATIO	−0.138	(0.03)	−0.107	(0.10)	−0.069	(0.34)	−0.048	(0.58)	0.010	(0.93)
R^2%	13.84		15.02		18.90		22.33		40.59	
Variable	*Z-Statistic*[d]	*p-value*	*Z-Statistic*	*p-value*	*Z-Statistic*	*p-value*	*Z-Statistic*	*p-value*	*Z-Statistic*	*p-value*
POSTFD	−2.736	(0.01)	−2.629	(0.01)	−2.108	(0.04)	−2.122	(0.03)	−3.559	(0.00)

(*continued on next page*)

TABLE 2 (continued)

---

[a] Variable definitions: $ACAR_{i,q,x}$ is the absolute cumulative abnormal return from x days before through 2 days after firm i's quarter q earnings announcement; $POSTFD_q$ = 1 if the observation is from the fourth quarter of 2000 or first or second quarters of 2001, and 0 otherwise; $RETVAR_{i,q}$ is the standard deviation of firm i's returns during the market model estimation period for the pre-FD quarter; $NEGCAR_{i,q}$ = 1 if the cumulative abnormal return from 64 days before through two days after the earnings announcement is negative, and 0 otherwise; $ABSCAR_{i,q}$ is the absolute value of the cumulative abnormal return from 64 days before through two days after the earnings announcement; $LOSS_{i,q}$ = 1 if firm i reports a loss in quarter q, and 0 otherwise; $NEGSPEC_{i,q}$ equals the absolute value of special items deflated by total assets if negative, and 0 otherwise; $BOND30_q$ is yield on 30-year U.S. Treasury bond as of the end of the quarter; $EPRATIO_{i,q}$ is firm i's quarter q earnings divided by firm i's price at the end of quarter q. Abnormal returns are prediction errors from the market model, estimated over the year ending the day before the start of the relevant fiscal quarter. We winsorize all continuous variables at the 99th percentile of the distributions of their absolute values.

[b] All p-values are two sided. p-values for means are from t-tests of the difference between the pre- and post-FD means. For medians, p-values are from Wilcoxon two-sample tests. p-values for regression parameters are based on ordinary least squares standard errors. P-values for Z-statistics are from the student's t-distribution. There are 9,510 observations.

[c] All correlations are significant at the 0.05 level or better (two-sided) except the correlations between (1) $NEGCAR_{i,q}$ and $NEGSPEC_{i,q}$, (2) $RETVAR_{i,q}$ and $BOND30_q$, and (3) $RETVAR_{i,q}$ and $POSTFD_q$, which are not significantly different from 0 at conventional levels.

[d] We estimate the equation for each of the 24 sets of seven consecutive quarters from the first quarter of 1993 through the second quarter of 2000, with $POSTFD_q$ equal to 1 for the last three quarters in each seven-quarter set and 0 for the first three quarters. The middle quarter is discarded. The Z-statistic is equal to $(\hat{a}_{1,x} - \bar{a}_{1,x})/\sigma_x$, where $\hat{a}_{1,x}$ is the estimate of $a_{1,x}$ from the true pre- and post-FD periods and $\bar{a}_{1,x}$ is the mean and σ_x the standard deviation of the 24 pseudo-$POSTFD_q$ coefficient estimates.

$POSTFD_q$ coefficient) from our actual pre- and post-FD quarters, and $\bar{a}_{1,x}$ is the mean and σ_x the standard deviation of the 24 pseudo-$POSTFD_q$ coefficient estimates. Panel C shows that, for all five accumulation windows, this Z-statistic is negative and statistically significant at the 0.05 level (two-sided) or better, suggesting the magnitude of the decrease in ACARs from pre- to post-FD is significantly larger than the average change in ACARs over the 1993–2000 period.

Although a number of the correlations among the control variables reported in Panel B of Table 2 are statistically significant, variance inflation factors (not tabulated) are less than 2.5, suggesting multicollinearity is inconsequential. The $RETVAR_{i,q}$ and $NEGCAR_{i,q}$ coefficients are positive as expected (although $NEGCAR_{i,q}$ is not significant for the longer −10,+2 and −30,+2 windows), suggesting firms with greater past price variability and firm-quarters with negative contemporaneous returns experience greater ACARs. The positive $ABSCAR_{i,q}$ coefficients indicate the ACARs for the various windows increase with the entire quarter's ACAR. Although not significant for every window, the $LOSS_{i,q}$ and $NEGSPEC_{i,q}$ coefficients are positive, suggesting that, conditional on other controls, ACARs are greater when firms report losses and negative special items. The negative $BOND30_q$ coefficients suggest lower ACARs in quarters with higher interest rates, consistent with higher future expected rates of return (Collins and Kothari 1989). The $EPRATIO_{i,q}$ coefficients are negative (although significant at the 0.10 level or better for the −1,+2 and −2,+2 windows only), suggesting short-window price responses are increasing in expected growth.

Our results are robust to various alternative specifications and design modifications. For example, we repeat all the univariate and regression analyses reported in Tables 1 and 2 separately for each of the three sets of like quarters in our sample period (i.e., the fourth, first, and second quarters) with generally similar results in each quarterly comparison. Our results are qualitatively similar when we estimate Equation (1) after explicitly controlling for both cross-sectional and time-series dependence.[22] We also estimate Equation (1) in seasonal first-difference form, where the intercept captures the effect of FD. Again, our inferences are unchanged. Finally, although we winsorize all our continuous variables, we also estimate Equation (1) after deleting observations where Cook's D is greater than 4/N (N = the number of observations) and, separately, after deleting observations with absolute studentized residuals greater than 2, with qualitatively similar inferences. Additionally, we estimate Equation (1) after replacing $ACAR_{i,q}$, $ABSCAR_{i,q}$, and $RETVAR_{i,q}$ with their natural logarithms, again, with qualitatively similar inferences. Thus, our inferences are unlikely driven by extreme values.

To summarize, our ACAR analyses suggest no decline in price efficiency prior to earnings announcements after the implementation of FD. On the contrary we find an *improvement* in the informational efficiency of stock prices post-FD that is systematic and statistically significant relative to random time-series variation in changes in ACARs. This improvement is also robust to controls for various determinants of ACARs. Overall, our evidence is consistent with improvement, rather than deterioration in firms' financial information environments post-FD.

---

[22] Since standard errors for the Z-statistics are from the time-series of parameter estimates, our Z-statistics are not subject to the influence of cross-sectional correlation (Bernard 1987). To control for serial dependence, we estimated Equation (1) using the first differences of the firm-specific pre- and post-FD averages (across the three quarters). In this specification, where the intercept captures the effect of FD, there is only one observation per firm. Finally, to control simultaneously for both cross-sectional and time-series dependence we estimate Equation (1) on the subset of firms with complete data in all three pre- and three post-FD quarters, using a generalized least squares procedure as described in Parks (1967). Our results are qualitatively unchanged under all these alternative specifications.

IV. EVIDENCE FROM ANALYSTS' FORECASTS

In this section we investigate the effect of FD on sell-side analysts' forecasts. The SEC issued FD to prevent managers from privately communicating material new information to securities professionals, such as analysts. Critics of FD suggest this will impair the accuracy of analysts' forecasts. However, analysts could respond to reduced information flows from managers by increasing their private information search, and firms may respond to additional limits on private communication by increasing public disclosures. Therefore, the effect of FD on analysts' performance is an empirical question.

Univariate Results

We consider two traditional measures of analysts' forecast performance: forecast error and dispersion. Following Brown et al. (1987), Ziebart (1990), Ajinkya et al. (1991), Lang and Lundholm (1996), and others, we measure forecast error as the absolute value of the difference between actual earnings per share and the mean analyst forecast, both as reported by First Call, as of the date of the most recently updated forecast prior to the earnings announcement. We measure forecast dispersion as the standard deviation of individual analysts' forecasts for a given firm, as of the date of the last forecast of quarterly earnings prior to the earnings announcement, since our objective is to examine changes in the information environment near the date of the earnings announcement. We scale both forecast error and dispersion by price at the end of the *control* (i.e., *pre*-FD) quarter, to prevent changes in our measures simply because of changing stock prices. Using price from the same day to deflate both our pre- and post-FD measures controls for cross-sectional scale differences without contaminating the results with deflator changes (Jacob et al. 1999). Finally, computing forecast dispersion requires that at least two analysts issue forecasts of the quarter's earnings. Thus, our analyses of forecast dispersion employ fewer observations than our analyses of forecast errors.

Table 3 reveals that mean and median absolute consensus forecast errors are significantly higher ($p = 0.02$ or better, two-sided) in the post-FD period than in the pre-FD period. Table 3 also presents Z-statistics for the significance of the differences in mean and median absolute consensus forecast errors, relative to their time-series variation, which we compute using a procedure analogous to that described in Section III. The Z-statistic for the difference in means is significant ($p = 0.02$, two-sided), although the Z-statistic is not significant for the difference in medians. However, more firms' absolute forecast errors increase than decrease after FD ($p < 0.01$, two-sided).

Table 3 also reveals that mean and median forecast dispersion are significantly higher post-FD, although the Z-statistic is significant only for the change in means. The proportion of firms with increased forecast dispersion significantly exceeds that with decreased dispersion ($p < 0.01$, two-sided). In summary, our univariate results provide some evidence that FD impaired analysts' forecast accuracy and increased dispersion.[23]

---

[23] Due, in part, to First Call's rounding of both actual and forecasted EPS to two decimal places, almost a fifth of the forecast errors are zero in both the pre- and post-FD periods. In cases where the forecast error for a firm is zero in both our pre- and matched post-FD quarters, the change in absolute forecast error is also zero. While the change in absolute forecast error is often positive or negative, these zero changes lie in the middle of the distribution, making the median of the within-firm changes in absolute forecast errors exactly zero. The median absolute forecast errors during both the pre- and the post-FD periods, however, are nonzero because the zero forecast errors cluster at the left tail of the distribution of forecast errors. Therefore, test results in Table 3 are based on changes in cross-sectional medians, and to maintain consistency, changes in cross-sectional means, rather than medians and means of the within-firm changes. However, tests based on means of the within-firm pre- to post-FD differences yield inferences similar to those drawn from Table 3.

TABLE 3
Absolute Analyst Forecast Error and Dispersion in Analysts' Forecasts Before and After FD[a]

	Absolute Forecast Error (n = 10,144)		Forecast Dispersion (n = 8,222)	
	Mean	*Median*	*Mean*	*Median*
Pre-FD	3.176[b]	1.049	1.353	0.588
Post-FD	4.096	1.127	1.596	0.682
Difference:				
Magnitude	0.920	0.078	0.243	0.094
p-value[c]	(0.00)	(0.02)	(0.00)	(0.00)
Z-Statistic[d]	2.350	0.579	1.981	1.349
p-value	(0.02)	(0.56)	(0.05)	(0.18)
% of Firm-Quarters with *Smaller* Forecast Metrics Post-FD	43.20		40.87	
% of Firm-Quarters with *Larger* Forecast Metrics Post-FD	45.17		55.85	
p-value	(0.00)		(0.00)	

---

[a] Variable definitions: Absolute forecast error is the absolute value of the difference between actual earnings and the mean of the individual analyst forecasts, both as reported by First Call. Dispersion is the standard deviation of the individual analysts' forecasts. We compute forecast dispersion using only firm-quarters where at least two analysts follow the firm. We scale absolute analyst forecast error and dispersion, for each pre-FD and corresponding post-FD quarter, by share price at the end of the pre-FD quarter, and winsorize scaled forecast error and dispersion at the 99th percentile of the distributions of their absolute values.

[b] For presentation purposes, we multiplied all means and medians in this table by 1,000.

[c] All p-values are two-sided. p-values for means are from t-tests of the difference between the pre- and post-FD means. For medians, p-values are from Wilcoxon two-sample tests. p-values for Z-statistics are from the student's t-distribution. p-values for difference in percentages are from binomial tests.

[d] We compute means and medians for each set of pseudo-pre- and pseudo-post-FD quarters in the 24 sets of seven consecutive quarters from the first quarter of 1993 through the second quarter of 2000. We label the first three quarters in each set pseudo-pre-FD quarters and the last three pseudo-post-FD quarters. The middle quarter is discarded. Z-statistics equal $(D - \overline{D})/\sigma$, where D is the post-FD mean or median forecast metric (i.e., absolute forecast error or dispersion) minus the pre-FD mean or median absolute forecast metric, and $\overline{D}$ and σ are the mean and standard deviation, respectively, of the 24 differences between the pseudo-post- and pseudo-pre-FD means or medians.

Regression Results

In this section we control for changes in the economic environment unrelated to FD that could be correlated with analysts' forecast performance. Specifically, we estimate the following model:

$$\begin{aligned}\text{Forecast Metric}_{i,q} = {} & b_0 + b_1\text{POSTFD}_q + b_2\text{SRWABSUE}_{i,q} + b_3\text{NEGUE}_{i,q} \\ & + b_4\text{LOSS}_{i,q} + b_5\text{NEGSPEC}_{i,q} + b_6\text{GDPSHOCK}_q \\ & + b_7\text{DAYS}_{i,q} + b_8\text{LAGMETRIC}_{i,q} + e_{i,q} \qquad (2)\end{aligned}$$

where Forecast Metric$_{i,q}$ is either the mean absolute forecast error or the standard deviation of individual analysts' forecasts. Again, our matched sample largely controls for stationary firm-specific factors. We therefore focus on controlling for factors affecting analysts' forecasting accuracy that likely changed from our pre- to post-FD periods, but are unrelated to FD.

We first include as controls several earnings characteristics that may affect forecasting accuracy. Since large changes in earnings are more difficult to forecast than small changes (e.g., Kross et al. 1990; Lang and Lundholm 1996), we include absolute seasonal random-walk unexpected earnings, deflated by pre-FD quarter-end stock price ($SRWABSUE_{i,q}$). Further, declining or negative earnings and earnings with negative special items may contain transitory components that analysts have difficulty anticipating. Therefore, we include the indicator variables $NEGUE_{i,q}$, which equals 1 if firm i's quarter q earnings are below earnings from the same quarter of the previous year, and 0 otherwise, and $LOSS_{i,q}$, which equals 1 if firm i's quarter q earnings are negative, and 0 otherwise. We also include $NEGSPEC_{i,q}$, which is defined as in Equation (1). Since analysts more likely experience difficulty forecasting earnings in periods of substantial economic change, we include the surprise in gross domestic product ($GDPSHOCK_q$), which we measure as the absolute value of the quarterly change in the seasonal growth rate in gross domestic product.[24] We include the log of the average number of days by which the forecast precedes the earnings announcement ($DAYS_{i,q}$), since later forecasts are more accurate than earlier forecasts (Kross et al. 1990; Clement 1999). Finally, since these variables likely do not fully capture firm-specific variation in our forecast metrics, we include each firm's forecast metric from the same quarter the year prior to our pre-FD quarters (i.e., $LAGMETRIC_{i,q} = LAGAFE_{i,q}$ or $LAGDISP_{i,q}$).[25]

Panel A of Table 4 presents descriptive statistics regarding our control variables. The mean of $SRWABSUE_{i,q}$ indicates that the current quarter's earnings differs, on average, from earnings four quarters prior by 1.1 percent of stock price, and the mean of $NEGUE_{i,q}$ indicates that the current quarter's earnings is below the same quarter's earnings from the previous year in 36.76 percent of our firm-quarters. During our sample period, the seasonal growth rate in gross domestic product changed by an average of about 0.95 percent. On average, 19.47 days (natural log is 2.969) elapse between the last forecast prior to the earnings announcement and the announcement. Over the three quarters 1 year prior to our three pre-FD quarters, analyst forecast errors averaged 0.3 percent of stock price and the average within-firm standard deviation of forecasts was 0.0013.

Panel A of Table 4 also presents pre- and post-FD means and medians for our control variables. The pre-FD means and medians are significantly different from the post-FD means and medians for all control variables (except $LAGAFE_{i,q}$ and $LAGDISP_{i,q}$, the two $LAGMETRIC_{i,q}$ variables, which, by construction, do not vary from our pre- to post-FD

---

[24] We obtained gross domestic product data from the Federal Reserve Bank of St. Louis web site: http://research.stlouisfed.org/fred/data/gdp.html.

[25] $LAGMETRIC_{i,q}$ differs for each of the three pre-FD quarters for each firm and these three values are repeated for the post-FD quarters. Since it does not vary from the pre- to the post-FD quarters, excluding it will not bias the coefficient on $POSTFD_q$, but will reduce the power of our tests and can, therefore, affect our inferences. We include it because we expect it to be an important explanator of forecast error and dispersion and we wish to minimize the risk of failing to reject the hypothesis that FD had no impact. We also estimate Equation (2) including several other cross-sectional variables suggested by prior research, including the firm's number of business segments, the firm's book-to-price ratio, institutional ownership, and the number of analysts following the firm (Dempsey 1989; Kross et al. 1990; Clement 1999; Chaney et al. 1999; Lim 2001). We also included the natural log of the firm's equity market value (Lang and Lundholm 1996) and the number of years since the firm listed on an exchange. Inclusion of these variables does not affect our inferences regarding the $POSTFD_q$ coefficient.

TABLE 4
Regression of Absolute Analyst Forecast Error or Dispersion on Post-FD Indicator and Control Variables

$$\textit{Forecast Metric}_{i,q} = b_0 + b_1 POSTFD_q + b_2 SRWABSUE_{i,q} + b_3 NEGUE_{i,q} + b_4 LOSS_{i,q} + b_5 NEGSPEC_{i,q} + b_6 GDPSHOCK_q + b_7 DAYS_{i,q} + b_8 LAGMETRIC_i + e_{i,q}$$

Panel A: Distributional Statistics

Control Variables[a]	Means Overall	Means Pre-FD	Means Post-FD	Means p-value[b]	Medians Overall	Medians Pre-FD	Medians Post-FD	Medians p-value
SRWABSUE	0.0110	0.0099	0.0120	(0.00)	0.0042	0.0039	0.0047	(0.00)
NEGUE	0.3676	0.2932	0.4421	(0.00)	0.0000	0.0000	0.0000	(0.00)
LOSS	0.1759	0.1572	0.1946	(0.00)	0.0000	0.0000	0.0000	(0.00)
NEGSPEC	0.0035	0.0024	0.0047	(0.00)	0.0000	0.0000	0.0000	(0.00)
GDPSHOCK	0.9546	0.5601	1.3492	(0.00)	0.6200	0.3960	1.6030	(0.00)
DAYS	2.9690	2.9248	3.0132	(0.00)	3.0910	3.0910	3.1355	(0.00)
LAGAFE	0.0030	0.0030	0.0030	NA	0.0009	0.0009	0.0009	NA
LAGDISP	0.0013	0.0013	0.0013	NA	0.0005	0.0005	0.0005	NA

Panel B: Correlations[c]

	SRWABSUE	NEGUE	LOSS	NEGSPEC	GDPSHOCK	DAYS	LAGAFE	LAGDISP
POSTFD	0.0565	0.1544	0.0491	0.0771	0.5940	0.0377	0.0000	0.0000
SRWABSUE		0.1476	0.2642	0.0865	0.0581	0.0305	0.3175	0.2768
NEGUE			0.3103	0.1111	0.0769	0.0015	0.0121	0.0042
LOSS				0.1410	0.0118	0.0573	0.2250	0.2359
NEGSPEC					0.0856	0.0225	0.0222	0.0170
GDPSHOCK						0.0722	−0.0203	−0.0231
DAYS							0.0582	0.0103
LAGAFE								0.4169

(continued on next page)

P17275

TABLE 4 (continued)

Panel C: Regression Results

	Absolute Forecast Error (n = 8,180)		Forecast Dispersion (n = 6,850)	
Variable	*Coefficient*	*p-value*	*Coefficient*	*p-value*
Intercept	−0.001	(0.00)	0.000	(0.00)
POSTFD[d]	0.246	(0.12)	−0.068	(0.24)
SRWABSUE	0.216	(0.00)	0.072	(0.00)
NEGUE	0.000	(0.05)	0.000	(0.00)
LOSS	0.003	(0.00)	0.001	(0.00)
NEGSPEC	0.024	(0.00)	0.004	(0.03)
GDPSHOCK	−0.000	(0.50)	0.000	(0.00)
DAYS	0.001	(0.00)	−0.000	(0.39)
LAGMETRIC	0.031	(0.00)	0.156	(0.00)
R^2%	40.65		36.95	
Variable	*Z-Statistic*[e]	*p-value*	*Z-Statistic*	*p-value*
POSTFD	1.082	(0.28)	1.219	(0.22)

(*continued on next page*)

TABLE 4 (continued)

---

[a] Variable definitions: $POSTFD_q$ = 1 if the observation is from the fourth quarter of 2000 or first or second quarters of 2001, and 0 otherwise; $SRWABSUE_{i,q}$ = absolute seasonal random-walk unexpected earnings; $NEGUE_{i,q}$, = 1 if earnings are below the same quarter's earnings from the previous year, and 0 otherwise; $LOSS_{i,q}$ = 1 if firm i reports a loss in quarter q, and 0 otherwise; $NEGSPEC_{i,q}$ equals the absolute value of special items deflated by total assets if the firm reports negative special items, and 0 otherwise; $GDPSHOCK_q$ is the absolute value of the quarterly change in the seasonal growth rate in gross domestic product; $DAYS_{i,q}$ = the log of the number of days the forecast precedes the earnings announcement; $LAGAFE_{i,q}$ = firm i's absolute forecast error from the same quarter the year prior to pre-FD quarter q; $LAGDISP_{i,q}$ = firm i's forecast dispersion from the same quarter the year prior to pre-FD quarter q; $LAGMETRIC_{i,q}$ = either $LAGAFE_{i,q}$ or $LAGDISP_{i,q}$. Absolute forecast error is the difference between actual earnings and the mean of the individual analyst forecasts. Forecast dispersion is the standard deviation of the individual analysts' forecasts. Absolute forecast error, forecast dispersion, and $SRWABSUE_{i,q}$, for each pre-FD and corresponding post-FD quarter, are divided by share price at the end of the pre-FD quarter. We winsorize all continuous variables at the 99th percentile of the distributions of their absolute values.

[b] All p-values are two sided. p-values for means are from t-tests of the difference between the pre- and post-FD means. For medians, p-values are from Wilcoxon two-sample tests. p-values for regression parameters are based on ordinary least squares standard errors. p-values for Z-statistics are from the student's t-distribution.

[c] All correlations are significant at the 0.05 level (two-sided) or better except for the correlation between $GDPSHOCK_q$ and $LAGAFE_{i,q}$, which is significant at the 0.06 level (two-sided), and the correlations between $GDPSHOCK_q$ and $LOSS_{i,q}$, $NEGUE_{i,q}$ and $DAYS_{i,q}$, $NEGUE_{i,q}$ and $LAGAFE_{i,q}$, $NEGUE_{i,q}$ and $LAGDISP_{i,q}$, $DAYS_{i,q}$ and $LAGDISP_{i,q}$, $POSTFD_q$ and $LAGAFE_{i,q}$, $LAGDISP_{i,q}$ and $NEGSPEC_{i,q}$, and $POSTFD_q$ and $LAGDISP_{i,q}$, which are not significant at conventional levels.

[d] For presentation purposes, we multiplied this coefficient by 1,000.

[e] We estimate the equation for each of the 24 sets of seven consecutive quarters from the first quarter of 1993 through the second quarter of 2000, with $POSTFD_q$ equal to 1 for the last three quarters in each seven-quarter set and 0 for the first three quarters. The middle quarter is discarded. The Z-statistic is equal to $(\hat{b}_1 - \bar{b}_1)/\sigma_{b_1}$ where $\hat{b}_1$ is the estimate of b_1 from the true pre- and post-FD periods, and $\bar{b}_1$ is the mean and σ_{b_1} the standard deviation of the 24 pseudo-$POSTFD_q$ coefficient estimates.

quarters). In particular, the difference between the pre- and post-FD means for $SRWABSUE_{i,q}$, $NEGUE_{i,q}$, $LOSS_{i,q}$, and $NEGSPEC_{i,q}$ suggest that, in our post-FD period, earnings had more time-series variability, were more often below four-quarter-ago earnings, were more likely negative, and contained more negative special items. Thus, to the extent these factors impair analysts' forecasting ability, it is important to control for their effects.

Panel B of Table 4 shows that the variables most correlated with $POSTFD_q$, are $GDPSHOCK_q$ and $NEGUE_{i,q}$, although $POSTFD_q$ is significantly correlated with all the control variables (except $LAGAFE_{i,q}$ and $LAGDISP_{i,q}$, which, by construction, do not vary from our pre- to post-FD quarters). Again, most of the correlations among the control variables are small in magnitude, but statistically significant ($p = 0.05$, two-sided).

We present results from estimating Equation (2) with absolute forecast errors as the dependent variable in Panel C of Table 4. The coefficient on $POSTFD_q$ is positive but significant at only the 0.12 level (two-sided). Further, the Z-statistic for the $POSTFD_q$ coefficient, which we construct analogously to the Z-statistics for Equation (1), is insignificant ($p = 0.28$, two-sided). We also estimate Equation (2) with forecast dispersion as the dependent variable. In this estimation, the $POSTFD_q$ coefficient is negative, but not statistically significant ($p = 0.24$, two-sided). Its Z-statistic is positive with a p-value (two-sided) of 0.22. Thus, our regression results suggest that, after controlling for various characteristics of the forecasting environment, there is no discernable change in the accuracy of analysts' forecasts or in their dispersion after FD's implementation.[26]

Coefficients on control variables suggest analysts have more trouble forecasting firm-quarters (1) with greater time-series forecast errors ($SRWABSUE_{i,q}$), (2) with declining earnings ($NEGUE_{i,q}$), (3) where firms report losses ($LOSS_{i,q}$), and (4) where firms report negative special items ($NEGSPEC_{i,q}$). Their forecasts are less accurate the longer the forecast horizon ($DAYS_{i,q}$) and they are able to reach less consensus in periods of significant economic change ($GDPSHOCK_q$). Past forecast error and past consensus ($LAGMETRIC_{i,q}$) are positively associated with current forecast error and current consensus, respectively. Despite the correlations mentioned above, none of the variance inflation factors suggest the presence of multicollinearity.

We repeat the analyses reported in Tables 3 and 4 separately for each of the three sets of like quarters in our sample period (i.e., the fourth, first, and second quarters), with generally similar results in each quarterly comparison. We estimate Equation (2) in seasonal first-difference form. Our inferences are unchanged. We also obtain qualitatively similar results when we estimate Equation (2) after explicitly controlling for both cross-sectional and time-series dependence, in similar manner to that in the case of the stock returns' analyses. Also, our estimations of Equation (2) after deleting observations where Cook's D is greater than $4/N$ ($N =$ the number of observations) and, separately, after deleting observations with absolute studentized residuals greater than 2 do not change our inferences. Our results are therefore unlikely driven by outlier effects.

In conclusion, our univariate evidence weakly suggests an increase in analysts' forecast errors and dispersion after FD. However, our regression analyses controlling for other likely determinants of analysts' forecast errors and dispersion (e.g., earnings variability, earnings declines, losses, negative special items, general economic changes, forecast timing, and other firm-specific factors) suggest no change in forecast errors or dispersion after FD. Thus, we conclude that the evidence does not support the contention that FD impaired analysts' forecasting ability.

---

[26] When we estimate Equation (2) in each size (defined as market value of equity) quartile, the results suggest analyst forecast errors *increased* post-FD for the *smallest* quartile and forecast dispersion *decreased* for the *largest* quartile. We leave further exploration of the differential impact of FD across firms to future research.

V. EVIDENCE FROM VOLUNTARY EARNINGS-RELATED DISCLOSURES

In principle, FD should prevent managers from privately communicating value-relevant information to select analysts. However, we find no deterioration in either the informational efficiency of stock prices or analysts' forecasting performance post-FD. We investigate whether firms are providing information through alternative channels, such as public disclosure.[27] This explanation is consistent with prior research suggesting firms are more likely to issue voluntary disclosures in the presence of greater information asymmetry (e.g., Ajinkya and Gift 1984; Diamond and Verrecchia 1991; Lang and Lundholm 1993). Accordingly, we examine whether the frequency of voluntary earnings-related disclosures increases after FD.

Univariate Results

Our voluntary disclosure data are from First Call's "Corporate Issued Guidance" database. We restrict our analyses to disclosures First Call classifies as "earnings-related." Our data include managers' forecasts of current period's and future periods' earnings, both quarterly and annual (although a few also pertain to revenue forecasts). We include point and range estimates, upper and lower bounds, and qualitative disclosures. We include all disclosures from the day after the previous quarter's earnings announcement to the day of the current quarter's earnings announcement for each of our six (three pre- and three post-FD) quarters.

Panel A of Table 5 presents descriptive information about the disclosures in our sample, both before and after FD. These disclosures more than doubled, from 1,160 in our pre-FD period to 2,981 in our post-FD period, consistent with firms responding to FD by increasing public disclosures. The change in composition is particularly striking. Prior to FD, approximately one-third (31.55 percent) of these disclosures were point estimates, where management provided a specific estimate, and approximately one-third (35.52 percent) were range estimates, where management provided a forecast of the range in which future earnings would fall. The remaining third included upper or lower bounds, purely qualitative descriptions, or other earnings-related information. After FD, however, nearly two-thirds (62.83 percent) of the voluntary earnings-related disclosures are range estimates, and less than one-quarter (21.67 percent) are point estimates. Thus, our data suggest that, while firms issued more forecasts of each specificity level after FD, they issued relatively more range and relatively fewer point estimates after FD than before.

Panel B of Table 5 presents the average number of voluntary earnings-related management disclosures per firm-quarter during the pre- and post-FD periods and also presents statistical tests of differences. The panel categorizes the disclosures into those about current (upcoming) vs. future earnings. We classify disclosures as about current earnings if the disclosure concerns the upcoming earnings report. We classify disclosures as about future earnings if the disclosure concerns future periods' earnings. The mean number of voluntary, earnings-related management disclosures per firm-quarter is less than 1, and the medians (not tabulated) are 0. However, the mean number of disclosures per firm-quarter increased significantly after FD, from 0.229 to 0.588. The mean number of disclosures about current earnings nearly doubled from 0.132 to 0.236, and the mean number of disclosures about future earnings more than tripled, from 0.097 to 0.352, after FD. All three differences are significant at better than the 0.01 level. The Z-statistics, computed analogously to those presented in earlier sections are significant at the 0.01 level, implying that the magnitudes of these increases in voluntary, earnings-related disclosures are unusual compared to their

---

[27] Another explanation is that analysts increased their private search activity in response to reduced information flows. Mohanram and Sunder (2002) find evidence consistent with this hypothesis.

TABLE 5
Voluntary Forward-Looking Earnings-Related Disclosures Firms Issued Before and After FD

Panel A: Types of Voluntary Disclosures

	Type of Disclosure										
	Point Estimates		*Range Estimates*		*Upper or Lower Bounds Only*		*Qualitative*		*Other Earnings Related*		
	Number	*Percent*	*Number*	*Percent*	*Number*	*Percent*	*Number*	*Percent*	*Number*	*Percent*	*Total*
Pre-FD	366	31.55	412	35.52	119	10.26	255	21.98	8	0.69	1,160
Post-FD	646	21.67	1,873	62.83	169	5.67	258	8.65	35	1.17	2,981
Change	290	−9.88	1,461	27.31	50	−4.59	3	−13.33	27	0.48	1,831

Panel B: Current and Future Quarters' Earnings-Related Disclosures

	Mean Disclosures per Firm-Quarter: All Firm-Quarters[a]			*Proportion of Firm-Quarters with at Least One Disclosure*			*Mean Disclosures per Firm-Quarter: Only Disclosing Firms*[c]		
	Current Earnings[b]	*Future Earnings*	*Total*	*Current Earnings*	*Future Earnings*	*Total*	*Current Earnings*	*Future Earnings*	*Total*
Pre-FD	0.132	0.097	0.229	11.24%	7.20%	15.04%	0.334	0.244	0.577
Post-FD	0.236	0.352	0.588	20.15%	21.27%	32.35%	0.595	0.889	1.484
Difference:									
Magnitude	0.104	0.256	0.359	8.91%	14.07%	17.31%	0.261	0.645	0.906
p-value[d]	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Z-Statistic[e]	2.862	8.741	5.746	2.377	7.150	4.693	1.402	5.202	3.231
p-value	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.16)	(0.00)	(0.00)
By Firm Differences:									
% of Firm-quarters with *Fewer* Disclosures Post-FD	7.69	4.51	8.62	7.26	4.02	7.26			
% of Firm-quarters with *More* Disclosures Post-FD	16.74	19.36	27.52	16.17	18.10	24.57			
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)			

(*continued on next page*)

TABLE 5 (continued)

---

[a] Mean Disclosures per Firm-Quarter: All Firm-Quarters is the total number of disclosures in a quarter, by all firms, divided by the number of sample firms, calculated separately for the pre- and post-FD periods.

[b] "Current Earnings" columns refer to voluntary disclosures about the upcoming earnings announcements. "Future Earnings" columns refer to voluntary disclosures about earnings for future periods. The "Total" column includes both types.

[c] Mean Disclosures per Firm-Quarter: Only Disclosing Firms is the total number of disclosures in a quarter, by just those firms making at least one disclosure in either the pre- or post-FD periods, divided by the total number of sample firms making at least one disclosure in either the pre- or post-FD periods.

[d] All p-values are two-sided. p-values for means are from t-tests of the difference between the pre- and post-FD means. For medians, p-values are from Wilcoxon two-sample tests. p-values for Z-statistics are from the student's t-distribution. p-values for difference in percentages are from binomial tests. There are 10,144 observations.

[e] We compute means and medians for each set of pseudo-pre- and pseudo-post-FD quarters in the 16 sets of seven consecutive quarters from the first quarter of 1995 through the second quarter of 2000. We label the first three quarters in each set pseudo-pre-FD quarters and the last three pseudo-post-FD quarters. The middle quarter is discarded. Z-statistics equal $(D - \overline{D})/\sigma$, where D is the post-FD mean minus the pre-FD mean, and $\overline{D}$ and σ are the mean and standard deviation, respectively, of the 16 differences between the pseudo-post- and pseudo-pre-FD means.

historical variation.[28] Finally, for both current and future earnings, significantly more firms increased than decreased their disclosures after FD.

The middle three columns in Panel B of Table 5 show that the proportion of firm-quarters containing at least one voluntary disclosure increases significantly from the pre- to the post-FD period. Across the three pre-FD quarters, only 15.04 percent of the firm-quarters contained at least one disclosure. After FD, firms increased disclosures about both current and future earnings such that 32.35 percent of the firm-quarters contained at least one disclosure.[29] The time-series Z-statistics are also positive and significant at better than the 0.02 level (two-sided). Overall, we find a marked increase in the number of firm-quarters with at least one voluntary, earnings-related disclosure after FD.

In the rightmost three columns, we track the mean number of disclosures per firm-quarter for only those firms making at least one voluntary disclosure in either our pre- or post-FD periods.[30] These columns reveal that disclosing firms increased their disclosure frequency after FD. The average number of disclosures per firm-quarter more than doubled from 0.577 prior to FD to 1.484 after FD. This increase is evident in disclosures about both current and future earnings. The average number of disclosures, per firm-quarter, about current earnings increased from 0.334 to 0.595, and the average number of disclosures, per firm-quarter, about future earnings more than tripled from 0.244 to 0.889. All are statistically significant, in standard tests, at the 0.01 level (two-sided) or better. The time-series Z-statistics are significant, however, only for disclosures about future earnings and for total disclosures.[31]

Regression Analyses

In this section, we report results from estimating an ordered logit model that controls for non-FD factors influencing managers' voluntary disclosure decisions. Again, we focus on factors that could have changed between the pre- and post-FD periods, rather than purely cross-sectional variables. The model we estimate is:

$$\begin{aligned}\mathrm{DISC}_{i,q} &= c_0 + c_1 + c_2\mathrm{POSTFD}_q + c_3\mathrm{ABSCAR}_{i,q} + c_4\mathrm{EPDEV}_{i,q}\\ &\quad + c_5\mathrm{EARLYAFE}_{i,q} + c_6\mathrm{EARLYDISP}_{i,q} + c_7\mathrm{DAMAGE}_{i,q}\\ &\quad + c_8\mathrm{NEGCAR}_{i,q} + c_9\mathrm{NEEDCAP}_{i,q} + e_{i,q} \qquad (3)\end{aligned}$$

where $\mathrm{DISC}_{i,q}$ equals 0 if the firm issues no earnings-related disclosures, 1 if the firm issues one disclosure, or 2 if the firm issues multiple disclosures during a quarter. POSTFD_q, as defined previously, is an indicator variable equaling 1 if the firm-quarter is post-FD, and 0 otherwise.

---

[28] Inspection of the data reveals that First Call did not collect this data comprehensively prior to 1995. The number of disclosures reported in the database stabilizes from approximately 1995 onward. Accordingly, we compute the Z-statistics reported in Tables 5 and 6 using only data from the first quarter of 1995 through the second quarter of 2000.

[29] Kile et al. (1998) report that between 39 and 53 percent of their sample make at least one voluntary disclosure per year. Approximately 45 percent of our sample firms make at least one voluntary disclosure in 2000 and approximately 55 percent make at least one voluntary disclosure in 2001, which is comparable to that reported by Kile et al.(1998). The frequencies we report in Table 5 are by quarter and are thus lower.

[30] Testing whether firms issuing disclosures in the pre-FD period issue more disclosures after FD is not practical because most firms do not issue disclosures in consecutive like quarters. Thus, if a firm discloses in a particular quarter, the odds are high that the firm will not disclose again in the like quarter next year. For this reason, we classify firms as "disclosing firms" if they disclose in either our pre- or post-FD quarters.

[31] Tests based on the means of the within-firm pre- to post-FD changes produce results very similar to those reported in Table 5.

Theoretical and empirical research suggests firms face incentives to issue disclosures to reduce information asymmetry between managers and various stakeholders. For example, the "expectation adjustment hypothesis" (Ajinkya and Gift 1984) suggests managers voluntarily disclose information when market expectations differ substantially from management's beliefs. We include several variables reflecting, albeit imperfectly, measures of information asymmetry between managers and investors in an attempt to control for the portion of information asymmetry not due to FD. Specifically, we include the absolute cumulative abnormal return over the quarter ($ABSCAR_{i,q}$), as larger price movements suggest greater information flows and, therefore, greater potential for managers to have information investors do not. We include the absolute deviation of the firms' earnings-price ratio from the industry median at the end of the fiscal quarter ($EPDEV_{i,q}$), since price-earnings multiples outside industry norms increase the likelihood investors' valuations differ from management's. We include the mean absolute analyst forecast error ($EARLYAFE_{i,q}$) and forecast dispersion ($EARLYDISP_{i,q}$), both measured on the day after the previous quarter's earnings announcement, since larger values of these variables suggest inaccurate and diverse initial expectations of earnings, increasing the likelihood that the market's expectations deviate from management's.

Research suggests managers have incentives to disclose when expected litigation costs are high (Skinner 1997). Following Skinner (1997), we measure potential damages from litigation as the dollar amount of investors' losses, abstracting from the change in the market index, from trading the firm's stock over the period under consideration. Specifically, $DAMAGE_{i,q}$ equals $-(MVE_{i,q})(MKTADJRET_{i,q})(1 - (1 - VOL_{i,q})^{N_{i,q}})$ where $MVE_{i,q}$ is the market value of equity 10 days before the first disclosure (10 days before the earnings announcement for firms without voluntary disclosures), $MKTADJRET_{i,q}$ is market-adjusted return, and $VOL_{i,q}$ is the average daily percentage of the firm's shares traded, both measured over the $N_{i,q}$ day period between the current and previous earnings announcement. As in Skinner (1997), we set negative values of $DAMAGE_{i,q}$ equal to \$1 million. We also include an indicator variable for negative returns ($NEGCAR_{i,q}$) as an additional proxy for litigation risk. Finally, to control for the *ex ante* need for capital, we include $NEEDCAP_{i,q}$, defined as operating cash flows less capital expenditures, scaled by current assets (Dechow et al. 1996).

Panel A of Table 6 reports descriptive data regarding the control variables. The mean quarterly absolute cumulative abnormal return ($ABSCAR_{i,q}$) is 0.2705 and cumulative abnormal returns are negative ($NEGCAR_{i,q}$) for 53.03 percent of the firm-quarters. Earnings-to-price ratios deviate, on average, from the industry median at the end of the quarter ($EPDEV_{i,q}$) by 0.55 percent of stock price. Forecast errors ($EARLYAFE_{i,q}$) and forecast dispersion ($EARLYDISP_{i,q}$) early in the quarter average 0.58 and 0.21 percent of stock price, respectively. $DAMAGE_{i,q}$ averages 97.39, indicating that average potential litigation damages, for the typical firm-quarter in our sample, are just shy of \$100 million. $NEEDCAP_{i,q}$ is negative, on average, suggesting the average firm in our sample needs external capital during our sample period (i.e., capital expenditures exceed operating cash flows).

The pre-FD means and medians are statistically different from their post-FD counterparts for all control variables. For our sample, absolute cumulative abnormal returns, the absolute difference between firm-specific and industry median earnings to price ratios, expected litigation damages, the percentage of sample firms with negative cumulative abnormal returns, and the need for external capital are all lower post-FD than pre-FD. Nonetheless, analysts' forecast errors and forecast dispersion, measured early in the quarter, are both higher in our post-FD period. Panel B of Table 6 reveals that most of the

TABLE 6
Ordered-Logit Regression of Voluntary Forward-Looking Earnings-Related Disclosure Occurrences on Post-FD Indicator and Control Variables

$$DISC_{i,q} = c_0 + c_1 + c_2POSTFD_q + c_3ABSCAR_{i,q} + c_4EPDEV_{i,q} + c_5EARLYAFE_{i,q} + c_6EARLYDISP_{i,q} + c_7DAMAGE_{i,q} + c_8NEGCAR_{i,q} + c_9NEEDCAP_{i,q} + e_{i,q} \quad (3)$$

Panel A: Distributional Statistics

Control	Means				Medians			
Variables[a]	*Overall*	*Pre-FD*	*Post-Fd*	*p-value*[b]	*Overall*	*Pre-FD*	*Post-FD*	*p-value*
ABSCAR	0.2705	0.2823	0.2587	(0.00)	0.1924	0.1975	0.1854	(0.02)
EPDEV	0.0055	0.0056	0.0054	(0.10)	0.0028	0.0030	0.0026	(0.00)
EARLYAFE	0.0058	0.0052	0.0063	(0.00)	0.0018	0.0017	0.0020	(0.00)
EARLYDISP	0.0021	0.0020	0.0022	(0.00)	0.0007	0.0007	0.0008	(0.00)
DAMAGE	97.3927	121.0256	73.7600	(0.00)	1.0000	1.0559	1.0000	(0.00)
NEGCAR	0.5303	0.5589	0.5017	(0.00)	1.0000	1.0000	1.0000	(0.00)
NEEDCAP	−0.0323	−0.0389	−0.0256	(0.09)	0.0050	0.0020	0.0082	(0.03)

Panel B: Pearson Correlations[c]

	ABSCAR	*EPDEV*	*EARLYAFE*	*EARLYDISP*	*DAMAGE*	*NEGCAR*	*NEEDCAP*
POSTFD	−0.0364	−0.0255	0.0523	0.0396	−0.0642	−0.0561	0.0185
ABSCAR		−0.0090	0.0700	0.0449	0.0304	−0.1425	−0.1187
EPDEV			0.1814	0.1468	−0.0405	−0.0392	0.0638
EARLYAFE				0.4417	−0.0680	0.0137	−0.1096
EARLYDISP					−0.0640	0.0074	−0.0952
DAMAGE						0.2248	0.0310
NEEDCAP							−0.0241

(continued on next page)

TABLE 6 (continued)

Panel C: Ordered-Logit Results

	Disclosures about:					
	Current Earnings[d]		*Future Earnings*		*Total*	
Variable	*Coefficient*	*p-value*	*Coefficient*	*p-value*	*Coefficient*	*p-value*
Intercept 1	−4.336	(0.00)	−3.408	(0.00)	−2.716	(0.00)
Intercept 2	−2.138	(0.00)	−2.337	(0.00)	−1.664	(0.00)
POSTFD	0.819	(0.00)	1.380	(0.00)	1.134	(0.00)
ABSCAR	−0.060	(0.60)	−0.403	(0.00)	−0.200	(0.06)
EPDEV	−0.446	(0.92)	10.158	(0.03)	5.359	(0.15)
EARLYAFE	28.844	(0.00)	−8.837	(0.03)	16.844	(0.00)
EARLYDISP	−46.943	(0.00)	−48.062	(0.00)	−52.264	(0.00)
DAMAGE[e]	0.557	(0.00)	0.690	(0.00)	0.672	(0.00)
NEGCAR	0.138	(0.04)	−0.103	(0.14)	0.007	(0.91)
NEEDCAP	0.527	(0.00)	0.588	(0.00)	0.589	(0.00)
Pseudo R^2 %	4.98		7.11		8.10	
Variable	*Z-Statistic*[f]	*p-value*	*Z-Statistic*	*p-value*	*Z-Statistic*	*p-value*
POSTFD	2.748	(0.00)	2.832	(0.00)	3.489	(0.00)

(continued on next page)

TABLE 6 (continued)

[a] Variable definitions: DISC$_{i,q}$ = 0 if the firm issues no voluntary earnings-related disclosures, 1 if the firm issues one disclosure, or 2 if the firm issues multiple disclosures during a quarter; POSTFD$_q$ = 1 if the observation is from the fourth quarter of 2000 or the first or second quarters of 2001, and 0 otherwise; ABSCAR$_{i,q}$ is the absolute value of the cumulative abnormal return from 64 days before through two days after the earnings announcement; EPDEV$_{i,q}$ = the absolute deviation of the firms' earnings-price ratio from the industry median at the end of the fiscal quarter; EARLYAFE$_{i,q}$ = the absolute mean forecast error as of the day after the previous quarter's earnings announcement; EARLYDISP$_{i,q}$ = the standard deviation in individual analysts' forecasts as of the same date as EARLYAFE$_{i,q}$; DAMAGE$_{i,q}$ = expected litigation damages computed as in Skinner (1997) in $ millions; NEGCAR$_{i,q}$ = 1 if the cumulative abnormal return from 64 days before through two days after the earnings announcement is negative, and 0 otherwise; NEEDCAP$_{i,q}$ = expected need for capital, computed as in Dechow et al. (1996). We winzorize all continuous variables at the 99th percentile of the distributions of their absolute values.

[b] All p-values are two sided. p-values for means are from t-tests of the difference between the pre- and post-FD means. For medians, p-values are from Wilcoxon two-sample tests. p-values for ordered-logit parameter estimates are based on the Wald test. p-values for Z-statistics are from the student's t-distribution. There are 7,064 observations.

[c] All correlations are significant at the 0.05 level (two-sided) or better except for the correlation between NEGCAR$_{i,q}$ and NEEDCAP$_{i,q}$, which is significant at the 0.06 level (two-sided), and the correlations between NEGCAR$_{i,q}$ and EARLYAFE$_{i,q}$, NEGCAR$_{i,q}$ and EARLYDISP$_{i,q}$, ABSCAR$_{i,q}$ and EPDEV$_{i,q}$, and POSTFD$_q$ and NEEDCAP$_{i,q}$, which are not significant at conventional levels.

[d] Current earnings columns refer to voluntary disclosures about the upcoming earnings announcements. Future earnings columns refer to voluntary disclosures about earnings for future periods. The total column includes both types.

[e] For presentation purposes, we multiplied this coefficient by 1,000.

[f] We estimate the equation for each of the 16 sets of seven consecutive quarters from the first quarter of 1995 through the second quarter of 2000, with POSTFD$_q$ equal to 1 for the last three quarters in each seven-quarter set and 0 for the first three quarters. The middle quarter is discarded. The Z statistic is equal to $(\hat{c}_2 - \bar{c}_2)/\sigma_{c_2}$, where $\hat{c}_2$ is the estimate of c_2, from the true pre- and post-FD periods and $\bar{c}_2$ is the mean and σ_{c_2} the standard deviation of the 16 pseudo-POSTFD$_q$ coefficient estimates.

correlations among these variables are statistically significant, although small in magnitude. The most notable exception is the relatively large correlation (0.44) between $EARLYAFE_{i,q}$ and $EARLYDISP_{i,q}$. The correlations between the post-FD indicator and the control variables are all less than 0.10.

We estimate Equation (3) separately for total disclosures, disclosures about current earnings, and disclosures about future earnings. Panel C of Table 6 displays the results. The $POSTFD_q$ coefficient is positive and significantly different from zero at better than the 0.01 level (two-sided) in all three regressions. The time series Z-statistic for the $POSTFD_q$ coefficient is also positive and significant at better than the 0.01 level. Thus, after controlling for various non-FD factors potentially influencing firms' voluntary disclosure decisions during our sample period, our evidence suggests firms increased their voluntary, forward-looking, earnings-related disclosures after FD.

The coefficients on our information asymmetry variables are neither consistently significant nor consistently of the expected sign. The coefficient on $ABSCAR_{i,q}$ is generally negative, contrary to expectations, but is not significant when disclosures about current earnings is the dependent variable. The $EPDEV_{i,q}$ coefficient is positive and significant only for forecasts of future quarters' earnings, which is sensible since price is a forward-looking variable, i.e., a deviation from the industry earnings-to-price ratio indicates expectations that the firm will perform differently than the industry in future quarters. The $EARLYAFE_{i,q}$ coefficient is positive and significant for current earnings forecasts alone, which is understandable since $EARLYAFE_{i,q}$, by design, measures poor expectations of current earnings. The coefficient on $EARLYDISP_{i,q}$ is negative and significant in all three regressions, inconsistent with the notion that managers make disclosures to reduce heterogeneity in analysts' expectations. However, this result may arise if the number of disclosures is correlated over time, and *past* disclosures reduce current analysts' forecast dispersion. As expected, the coefficient on our litigation risk proxy, $DAMAGE_{i,q}$, is positive and significant in all models. The coefficient on $NEGCAR_{i,q}$ is positive and significant when we estimate Equation (3) for disclosures about current earnings, but is not significant otherwise. The coefficient on $NEEDCAP_{i,q}$ is consistently positive and significant.

We repeat all the univariate and the ordered logit analyses reported in Tables 5 and 6 separately for each of the three sets of like quarters in our sample period with qualitatively similar results in each quarterly comparison. We also estimated Equation (3) including several other variables suggested by prior research, including institutional ownership, number of employees scaled by total assets, number of analysts following the firm, and the natural log of the market value of equity (see Lang and Lundholm 1993, 2002; Healy et al. 1999; Miller and Piotroski 2000; Bushee and Noe 2001; Chen 2002). We also included the book-to-price ratio as another measure of uncertainty about firm value (Barth et al. 2001) and trading volume, as a percentage of outstanding shares, as a measure of heterogeneous beliefs. Results regarding the $POSTFD_q$ coefficient are virtually identical. Estimating Equation (3), after replacing $ABSCAR_{i,q}$ and $DAMAGE_{i,q}$ with their natural logs, does not change our inferences.

In summary, our evidence suggests firms increased their voluntary disclosures of forward-looking earnings information after implementation of FD. Our univariate analyses also suggest the proportion of firms making disclosures increased after FD and that disclosing firms increased their disclosure frequency after FD. Our results are robust to controls for alternative non-FD motivations for voluntary disclosures, and are unusual relative to time-series variation.

VI. CONCLUSION

Critics of FD suggest private communication to analysts is an important source of material information to the capital markets. They argue that, by essentially prohibiting private communication between firms and analysts, FD will impair the average level of information in the capital markets, resulting in less accurate expectations of firm performance and greater price shocks when performance is revealed. In this paper, we analyze the effect of FD on the financial information environment prior to a key source of information to the capital markets: quarterly earnings announcements. We find absolute cumulative abnormal returns prior to earnings announcements are significantly smaller post-FD than pre-FD. These decreases are significant compared to random variation in absolute cumulative abnormal returns over time and robust to controls for other factors affecting the earnings-returns relation and suggest stock prices incorporate earnings-related information earlier after FD's implementation. Simple univariate tests on mean and median absolute forecast errors and forecast dispersion suggest analysts' forecasting ability declined after FD's implementation. However, these impairments are not statistically significant relative to the historical time-series of changes in forecast errors and dispersion, and we find no significant reduction in analysts' performance after FD when we control for non-FD determinants of analysts' forecast accuracy and dispersion. Finally, we find a marked increase in the frequency of voluntary forward-looking disclosures by firms after FD became effective, even after controlling for other factors explaining voluntary disclosure.

In summary, other than mixed and weak univariate evidence on analysts' forecast performance, we find no reliable evidence of significant deterioration in the information environment prior to earnings announcements, after implementation of FD. On the contrary, some of our analyses suggest an improvement. For example, we find improved information efficiency of stock prices and a substantial increase in the frequency with which firms voluntarily disclose forward-looking earnings-related information. Our evidence that firms issue more voluntary, earnings-related disclosures is consistent with firms substituting voluntary, public disclosures to at least partially offset reductions in information flows through analysts.

While we have attempted to control for other non-FD factors that could affect our inferences, since our event (implementation of FD) occurs simultaneously for all firms, we can never completely rule out the possibility that our results are attributable to some other unknown contemporaneous economic event unrelated to FD. Also, our study documents *average* effects. We leave to future research examinations of cross-sectional differences in the effect of Regulation FD. Finally, our study investigates the immediate short-term effects of FD. While evidence on the immediate effect of FD is of interest from a policy perspective, the long-term effects of FD are unknown. Thus, although our results provide interesting early evidence on the short-term effect of Regulation FD, the study's results should be interpreted with caution.

REFERENCES

Ajinkya, B., and M. Gift. 1984. Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research* 22 (Autumn): 425–444.

———, R. Atiase, and M. Gift. 1991. Volume of trading and the dispersion in financial analysts' forecasts. *The Accounting Review* 66 (April): 389–401.

Association for Investment Management and Research (AIMR). 2001. FD e-Survey summary. Retrieved June 2001 from http://www.aimr.com/pressroom/01releases/regfd_surveysum.htm.

Atiase, R. K. 1985. Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research* 23 (Spring): 21–36.

Ball, R., and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (Autumn): 159–178.

———, and E. Bartov. 1996. How naive is the stock market's use of accounting information? *Journal of Accounting and Economics* 21 (June): 319–337.

Bamber, L. S., T. E. Christensen, and K. M. Gaver. 2000. Do we really "know" what we think we know? A case study of seminal research and its subsequent overgeneralization. *Accounting, Organizations and Society* 25 (February): 103–129.

Barth, M., R. Kasznik, and M. McNichols. 2001. Analyst coverage and intangible assets. *Journal of Accounting Research* 39 (June): 1–34.

Beaver, W. H. 1968. The information content of annual earnings announcements. *Journal of Accounting Research* 6 (Supplement): 67–92.

———, R. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting and Economics* 2 (March): 3–28.

Bernard, V. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* 25 (Spring): 1–48.

———, and J. Thomas. 1989. Post-earnings-announcement drift: Delayed price response or risk premium. *Journal of Accounting Research* 27 (Supplement): 1–36.

———, and ———. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13 (December): 305–340.

Black, F. 1976. Studies of stock price volatility changes. In *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Section*, 177–181.

Brown, L. D., R. L. Hagerman, P. A. Griffin, and M. E. Zmijewski. 1987. Security analyst superiority relative to univariate time series models in forecasting quarterly earnings. *Journal of Accounting and Economics* 9 (April): 61–88.

Bushee, B., and C. Noe. 2001. Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research* 38 (Supplement): 207–246.

Chaney, P., C. Hogan, and D. Jeter. 1999. The effect of reporting restructuring charges on analysts' forecast revisions and errors. *Journal of Accounting and Economics* 27 (June): 261–284.

Chen, S. 2002. Managerial guidance of market expectations: Incentives and effects. Working paper, University of Southern California, Los Angeles.

Cheon, Y. S., T. E. Christensen, and L. S. Bamber. 2001. Factors associated with differences in the magnitude of abnormal returns around NYSE vs. NASDAQ firms' earnings announcements. *Journal of Business Finance and Accounting* 28 (November): 1073–1108.

Christie, A. A. 1982. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics* 10 (December): 407–432.

Clement, M. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27 (June): 285–303.

Collins, D., and S. P. Kothari. 1989. An analysis of the intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics* 11 (July): 143–181.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1996. Causes and consequences of earnings manipulations: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (Spring): 1–36.

Dempsey, S. J. 1989. Predisclosure information search incentives, analyst following, and earnings announcement price response. *The Accounting Review* 64 (October): 748–757.

Diamond, D. W., and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *Journal of Finance* 46 (September): 1325–1359.

Francis, J., D. Hanna, and L. Vincent. 1996. Causes and effects of discretionary asset writeoffs. *Journal of Accounting Research* 34 (Supplement): 117–134.

———, K. Schipper, and L. Vincent. 2001. Earnings announcements and competing information. Working paper, Duke University, Durham, NC.

———, ———, and ———. 2002. Expanded disclosures and the increased usefulness of earnings announcements. *The Accounting Review* 77 (July): 515–546.

Freeman, R. N. 1987. The association between accounting earnings and security returns for large and small firms. *Journal of Accounting and Economics* 9 (July): 195–228.

———, and S. Tse. 1989. The multi-period information content of accounting earnings: Confirmations and contradictions of previous earnings reports. *Journal of Accounting Research* 27 (Supplement): 49–79.

Grant, E. B. 1980. Market implications of differential amounts of interim information. *Journal of Accounting Research* (Spring): 255–268.

Greene, W. H. 1997. *Econometric Analysis.* Upper Saddle River, NJ: Prentice Hall.

Hassett, K. A. 2000. Outlaw selective disclosure?—No, the more information the better. *Wall Street Journal* (August 10): A18.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 29 (September): 125–153.

Healy, P., A. Hutton, and K. Palepu. 1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research* 16 (Fall): 485–520.

Holthausen, R. W., and R. E. Verrecchia. 1988. The effect of sequential information releases on the variance of price changes in an intertemporal multi-asset market. *Journal of Accounting Research* 26 (Spring): 82–106.

Jacob, J., T. Lys, and M. Neale. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28 (November): 51–82.

———, ———, and J. Sabino. 2000. Autocorrelation structure of forecast errors from time-series models: Alternative assessments of the causes of post-earnings announcement drift. *Journal of Accounting and Economics* 28 (December): 329–358.

Kile, C., G. Pownall, and G. Waymire. 1998. How frequently do managers disclose prospective earnings information? *Journal of Financial Statement Analysis* 3 (Spring): 5–16.

Kim, M. S., and W. Kross. 2002. Differences between market responses to earnings announcements in the 1990s vs. 1960s. Working paper, University of Missouri, Columbia, MO.

Kross, W., and D. Schroeder. 1989. Firm prominence and the differential information content of quarterly earnings announcements. *Journal of Business, Finance, and Accounting* 16 (Spring): 55–74.

———, B. Ro, and D. Schroeder. 1990. Earnings expectations: The analysts' information advantage. *The Accounting Review* 65 (April): 461–476.

Landsman, W., and E. Maydew. 2002. Has the information content of annual earnings announcements declined in the past three decades? *Journal of Accounting Research* 40 (June): 797–808.

Lang, M. H., and R. J. Lundholm. 1993. Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research* 31 (Autumn): 246–271.

———, and ———. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71 (October): 467–492.

———, and ———. 2000. Voluntary disclosure and equity offerings: Reducing information asymmetry or hyping the stock? *Contemporary Accounting Research* 17 (Winter): 623–662.

Lim, T. 2001. Rationality and analysts' forecast bias. *The Journal of Finance* 56 (February): 369–385.

Lo, A. W., and A. C. MacKinlay. 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies* 1 (Spring): 41–66.

Mendenhall, R., and W. Nichols. 1988. Bad news and differential market reactions to announcements of earlier-quarters vs. fourth-quarter earnings. *Journal of Accounting Research* 26 (Spring): 63–90.

———, and D. Fehrs. 1999. Option listing and the stock price response to earnings announcements. *Journal of Accounting and Economics* 27 (February): 57–87.

Miller, G., and J. Piotroski. 2000. Forward looking earnings statements: Determinants and market response. Working paper, Harvard University, Cambridge, MA.

Mohanram, P., and S. Sunder. 2002. Has regulation fair disclosure affected financial analysts' ability to forecast earnings? Working paper, New York University, New York, NY.

Morse, D. 1981. Price and trading volume reaction surrounding earnings announcements: A closer examination. *Journal of Accounting Research* 19 (Autumn): 374–398.

Opdyke, J. D., and A. Lucchetti. 2000. Mum's the word in wake of disclosure rule. *Wall Street Journal* (August 16): C1.

Parks, R. W. 1967. Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. *Journal of the American Statistical Association* 62 (June): 500–509.

Penman, S. 1996. The articulation of price-earnings ratios and market-to-book ratios and the evaluation of growth. *Journal of Accounting Research* 34 (Autumn): 235–259.

Pulliam, S. 1999. Abercrombie & Fitch ignites controversy over possible leak of sluggish sales data. *Wall Street Journal* (October 14): C1.

Securities Industry Association (SIA). 2001. Costs and benefits of regulation fair disclosure. New York, NY: SIA.

Shane, P., N. Soderstrom, and S. W. Yoon. 2001. Earnings and price discovery in the post-reg. FD information environment: A preliminary analysis. Working paper, University of Colorado at Boulder.

Skinner, D. 1997. Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics* 23 (November): 249–282.

Weber, J. 2000. Full disclosure for all the web allows corporations to share information quickly and cheaply. *Business Week* (September 18): 106.

Williams, M., and R. McGough. 2000. Intel's jolt shows shifts in market's dynamics—New disclosure rules mean more legwork ahead for analysts. *Wall Street Journal* (September 25): C1.

Wild, J. J., L. A. Bernstein, and K. R. Subramanyam. 2001. *Financial Statement Analysis*. Seventh edition. New York, NY: McGraw-Hill/Irwin.

Ziebart, D. 1990. The association between consensus of beliefs and trading activity surrounding earnings announcements. *The Accounting Review* 65 (April): 477–488.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 39–69

Social Behaviors, Enforcement, and Tax Compliance Dynamics

Jon S. Davis
University of Wisconsin–Madison
Gary Hecht
Jon D. Perkins
University of Illinois at Urbana–Champaign

ABSTRACT: We analyze the effect of social norms and enforcement on the dynamics of taxpayer compliance. Specifically, we develop two models to evaluate the movement between classes of compliant and noncompliant taxpayers. Our analysis suggests that the effect on compliance of changing enforcement levels depends on whether the taxpayer population is initially compliant or noncompliant. Compliant populations are insensitive to changes in enforcement policies until enforcement becomes sufficiently lax, when we observe a sudden shift to high levels of noncompliance in equilibrium. In contrast, relatively noncompliant populations respond to increased enforcement by gradually increasing compliance. Then, when enforcement becomes sufficiently harsh, we find a sudden shift in equilibrium to very high levels of compliance. After the taxpayer population shifts from compliance to noncompliance, or vice versa, our models predict that returning to the previous enforcement policy will not cause the population to return to its previous state. On the whole, our models' results help explain why taxpayer compliance varies across time and across geographic regions, even under similar enforcement regimes.

Keywords: *regulation; tax compliance; agent-based modeling; nonlinear dynamics.*

Data Availability: *Contact the authors.*

The authors are deeply indebted to Ralph Wojowitcz for his research assistance. We also acknowledge the valuable comments and insights provided by Richard Gaylord, Young Kwon, Norman Solomon, Dan Stone, Robert Yetman and by workshop participants at the 2000 American Accounting Association Annual Meeting, Washington University in St. Louis, University of Illinois at Urbana–Champaign, and University of Wisconsin–Madison.

Submitted May 2000
Accepted August 2002

Morals cannot be legislated but behavior can be regulated.

—Martin Luther King, Jr.

I. INTRODUCTION

Enforcement and the behavior of others affect taxpayer compliance. IRS Commissioner Charles Rossotti recently stated that "the slide in audits has raised concern that honesty could suffer as fears of policing decline. If taxpayers begin to believe that others are cheating, the temptations to shave their own tax burdens may become irresistible" (Weisman 2001, A1). Rossotti's comment recognizes that tax enforcement affects social behaviors, and that these behaviors can, in turn, affect taxpayers' compliance decisions. Prior tax research concludes that individuals who know evading taxpayers are less likely to comply themselves (Vogel 1974; Spicer and Lundstedt 1976; Scott and Grasmick 1981; Grasmick and Scott 1982).[1] Likewise, the probability that a taxpayer will evade increases when the taxpayer suspects that his acquaintances are evading (Westat, Inc. 1980; Geeroms and Wilmots 1985). At the societal level, survey evidence suggests that differences in compliance across countries and geographic regions with similar or nearly identical tax and enforcement regimes stem from differences in social norms.[2] In an experiment, Alm et al. (1995) show that compliance differs across subjects from different countries. Although survey, archival, and experimental research support the contention that social norms and acquaintances' taxpaying behavior affect tax compliance, no theoretical foundation that exists suggests the form that these effects might take (Alm 1991).

We address this void in the literature by developing two models in order to provide a theoretical analysis of the effects of others' behavior, enforcement, and social norms on the dynamics of taxpayer compliance. The first, an aggregate-level model, uses a system of nonlinear differential equations to characterize the flows among three classes of taxpayers who differ in their propensities to evade: honest taxpayers, susceptible taxpayers, and evaders.[3] The flows represent the movements of taxpayers among the classes resulting from the aggregate outcomes of individual taxpayers' decisions. As the number of tax evaders increases, honest people become susceptible to evasion (honest taxpayers become increasingly likely to know evaders and to be influenced by their behavior). Similarly, the greater the proportion of honest taxpayers in the population, the greater the pressure on evaders to become honest. The level of enforcement affects the movement of taxpayers both from the susceptible class to the evading class, and from the evading class to the honest class.

To maintain tractability, we make a number of simplifying assumptions regarding aggregate behavior and we do not explicitly account for the behavior of individual taxpayers.

---

[1] Hite (1988) is an exception in this line of research. In a role-playing experiment using prospective jurors, she finds that peer influence has no significant effect on tax compliance.

[2] We obtained survey evidence regarding taxpayer behavior in Spain, Sweden, and the United Kingdom from de Juan et al. (1994), Vogel (1974), and Lewis (1979), respectively. For evidence regarding the United States, see Westat, Inc. (1980) and Yankelovich, Skelly, and White, Inc. (1984). These studies' results suggest that cross-cultural differences in the social acceptability of evasion may partly account for differences in observed levels of compliance. In addition, these studies found that a number of cultural variables that mediate the effect of norms on behavior (e.g., social cohesion, alienation, etc.) were significantly related to compliance behavior. Evidence also suggests that regional differences in compliance within countries may be attributable to local differences in social norms (e.g., Witte and Woodbury [1985] note an apparent effect within the United States).

[3] Mathematical biologists commonly use our approach to model epidemics (see Kermack and McKendrick 1927; Murray 1989, 611–618). Our model differs considerably from those in epidemiology, however, in that we carefully constructed the flows in our model to represent the social psychology of tax evasion. Researchers have used this class of models to investigate a number of topics in the social sciences, ranging from a model of social revolution to the spread of illegal drug use (Epstein 1997, 69–93).

To address these limitations, we create a second, complementary, model that characterizes individual taxpayers' compliance decisions. This micro-level model uses a relatively new method, agent-based modeling, which relaxes unrealistic assumptions of agent rationality and agent homogeneity, avoids the use of "representative individuals,"[4] and facilitates the theoretical representation of agent interaction and dynamic behavior.[5] The model does not characterize aggregate behavior explicitly. Instead, the aggregate behavior emerges endogenously and dynamically from the interactions among heterogeneous agents.

In our agent-based model, we create a society of taxpayers who initially possess limited knowledge about the level of enforcement and heterogeneous reporting rules that depend on perceived enforcement severity, social norms, and acquaintances' tax compliance behavior. Our taxpayers develop and periodically update their beliefs regarding social norms and enforcement severity. Individual taxpayer behavior evolves over time, partly in response to other taxpayers' behavior.

The results of our two models contribute to an understanding of how social forces and enforcement affect compliance. Our aggregate model suggests that taxpayers tend toward one or two equilibrium states over time. The possibility of multiple equilibria is consistent with the observation that tax compliance varies over time and across geographic regions, even under similar enforcement policies. Perhaps the most important prediction is that small changes in enforcement can lead to sudden and large changes in equilibrium states from mostly compliant to mostly evading or vice versa. For example, we predict that reducing enforcement in a compliant population will lead to a gradual increase in evasion, but at some point, a sudden proliferation of tax evasion will occur, effectively leading to an evasion epidemic. Our analysis further suggests that simply returning to the previous level of enforcement will not be sufficient to stop the epidemic. Our agent-based analysis supports the aggregate model's predictions.

We also make a methodological contribution to the accounting literature by introducing two complementary modeling approaches. In contrast to standard neoclassical economics, our aggregate model builds on a social-psychological foundation and explicitly accounts for system dynamics. Our agent-based model also permits an investigation of system dynamics and allows us to evaluate the robustness of the aggregate model's predictions. This approach may be useful in other areas of accounting research, such as the dynamics of reputation in public accounting, earnings management, the adoption and spread of managerial accounting practices, and the effect of other audit firms' behaviors on auditor independence.

The remainder of the paper begins by describing the assumptions underlying our aggregate-level model. We then analyze the model and develop comparative statics. Next, we introduce the agent-based model and describe a series of numerical experiments examining its comparative statics properties. Finally, we discuss the implications arising from our results, the limitations of this research, and possible future avenues of inquiry.

II. AGGREGATE MODEL ASSUMPTIONS

We model the population of taxpayers in an undefined geographic region; our results could apply to the population of a city, a state, an administrative region of the tax agency, or a country. For simplicity, we assume that the population is constant and consists of three discrete groups of taxpayers. The first group, x_{1t}, is the proportion of ***honest*** taxpayers in the population.[6] Honest taxpayers comply at time period t. These taxpayers do not consider evasion. They are either habitually compliant (Erard and Feinstein 1994) or they are recent evaders who have become honest as a result of enforcement efforts or social norms. The second group consists of taxpayers who are dissatisfied with the tax system (perhaps as a result of seeing others evade without being punished). These taxpayers are not actively evading, but they might if the perceived benefits of doing so exceed the perceived costs. For this group, evasion is an option, and so we classify them as *susceptible*. The proportion of susceptible taxpayers in the population at time period t is x_{2t}. The third group, x_{3t}, represents the proportion of *evading* taxpayers in the population at time period t. Whether a taxpayer continues to evade depends on both enforcement and the effect of social norms.

---

[4] Kirman (1992) critiques the use of representative agents in analytical models.

[5] Agent-based models are not related to agency theory. Epstein and Axtell (1996) provide an excellent review and a general introduction to agent-based models in the social sciences.

Because the entire population is made up of these three classes of taxpayers:

$$x_1 + x_2 + x_3 = 1. \tag{1}$$

We further assume that each proportion is initially positive:

$$x_{10} > 0; \quad x_{20} > 0; \quad x_{30} > 0. \tag{2}$$

We consider any population that satisfies these constraints as valid.

We next describe and model taxpayers' movements among the three classes resulting from the behavior of others, enforcement, and social norms. The model is a system of differential equations.

The Effect of Others' Behavior

Social psychologists have offered several explanations why knowing a tax evader might cause an honest taxpayer to consider evasion. For example, Lerner (1998) suggests that people need to believe the world is just. Consequently, when people observe an unjust event, they may cope by punishing the harm-doer, compensating or blaming the victim, or denying the injustice by reasoning that justice will prevail in the next life. In some cases, such as tax evasion, one might seek justice by engaging in the activity oneself (Spicer and Becker 1980; Tyler 1990).

An alternative explanation for why knowing a tax evader might cause honest taxpayers to consider evasion is that observing others' behavior can affect one's own internalized moral standards. Cooter (1998) argues that people prefer conformity to behavioral standards. When someone violates a standard, they incur a psychological cost—guilt—whether or not others discover the behavioral violation. However, if others of perceived high moral character violate a law, then one's behavioral standard may change. For example, Kaplan and Reckers (1985) provide experimental evidence that subjects were more likely to evade taxes when they observed a taxpayer of perceived high moral character evading.

Consistent with these observations and prior research, in our model as an individual encounters noncompliant taxpayers, he is more likely to consider evasion. We represent this relationship via the following system of differential equations:

---

[6] Because we focus on population proportions, our model provides limited insight into the dollar amount of taxes evaded. Evaders can choose activities that permit greater opportunities for evasion, leading to larger dollar amounts evaded than the population proportion would suggest. Other factors may reduce the amount of taxes evaded (e.g., third-party reporting requirements may restrict evasion).

$$\dot{x}_1 = \frac{dx_1}{dt} = -rx_1x_3$$
$$\dot{x}_2 = \frac{dx_2}{dt} = rx_1x_3 \quad (3)$$

where r represents the "infection rate." For simplicity, we model the rate of change of the honest taxpayer class, x_1, as a proportion of the product of honest taxpayers, x_1, and evading taxpayers, x_3. This is a standard simplification in epidemic models (e.g., Waltman 1974, 2) and implies uniform mixing of the groups and no delay before becoming susceptible.[7]

Modeling the Enforcement Regime

We assume that susceptible and evading taxpayers base their tax-reporting decisions partly on a psychological calculus that trades off the perceived costs and benefits of evasion. A taxpayer's view of these costs and benefits depends on enforcement (e.g., penalties, audit likelihood, audit costs imposed on the taxpayer, the public image of the tax agency, etc.), tax rates and their progressivity, and psychological factors, including framing effects from under- or overwithholding.[8]

We model two effects of enforcement on individual tax compliance behavior. First, we model a flow from the evader class to the honest class. For simplicity, we assume that evaders become compliant after they are audited (akin to a change in driving behavior after receiving a traffic ticket)[9] or when their perceptions regarding the costs and benefits of evasion change, either through experience or changing economic conditions.[10] We also assume that this flow from the evader to the honest class as a result of *enforcement* is proportional (at rate β) to the size of the evader class. This assumption is consistent with the tax agency behaving strategically: as the number of evaders increases, the tax authority increases enforcement efforts and vice versa.

Enforcement also affects "susceptible" (but not "honest") taxpayers' behavior through its effect on the perceived costs of evasion. Recall that susceptible taxpayers' compliance decisions are based on a cost-benefit analysis. We assume that some susceptible taxpayers will perceive that the benefits of evasion exceed the costs of evasion in each period. These taxpayers will evade, creating a flow from x_{2t} to x_{3t}. We assume that this flow is proportional to the size of the susceptible class (at rate α, where α is between 0 and 1). As enforcement increases, the cost of evasion increases, reducing the flow to the evading class. Similarly, decreases in enforcement will increase the flow to the evading class. When we add both enforcement-regime effects to the system of differential equations in Equation (3), they become:

---

[7] We test the robustness of the uniform mixing assumption and other assumptions regarding the flows among taxpayer classes using our agent-based model.

[8] For example, underwithheld taxpayers tend to perceive payments to the tax agency as losses while overwithheld taxpayers tend to perceive refunds from the tax agency as gains. Because individuals are risk seeking in perceived losses and risk averse in perceived gains (Kahneman and Tversky 1979), under- or overwithholding can affect tax compliance (e.g., Schepanski and Kelsey 1990; White et al. 1993).

[9] Psychological explanations for post-audit compliance include availability (Tversky and Kahneman 1973) and vividness (Nisbett et al. 1976; Plous 1993). This research generally finds that taxpayers judge an audit as more probable if it is vivid or more easily recalled. In addition, taxpayers may experience guilt or stigma if the tax agency discovers their evasion. Combined with availability, it is reasonable to expect such taxpayers to report honestly, at least for a time.

[10] Some taxpayers might flow from the evader class back to the susceptible class. That is, some evaders might not be rehabilitated when they are audited, remaining susceptible rather than becoming honest. Similarly, one might add a flow from the susceptible class back to the honest class to allow for spontaneous recovery from susceptibility. Adding these flows to our model complicates the analysis but the results do not differ significantly from those reported here. Specifically, these alternative models have stable honest and mixed equilibria, they meet our plausibility criteria, and the comparative statics results are qualitatively the same as those reported below.

$$\begin{aligned}\dot{x}_1 &= \beta x_3 - r x_1 x_3\\ \dot{x}_2 &= -\alpha x_2 + r x_1 x_3\\ \dot{x}_3 &= \alpha x_2 - \beta x_3\end{aligned} \tag{4}$$

Incorporating the Effect of Social Norms

The role of social norms is distinct from the role of others' compliance behavior. A social norm requires consensus about the esteem-worthiness of a behavior and some risk that a member of society will detect behavior that deviates from the norm. Society-at-large imposes sanctions on individuals who are caught violating a social norm.[11] In contrast, as noted above, the effect of others' behavior does not require consensus about the esteem-worthiness of an action, and one's response to the action is independent of detection (e.g., one feels guilty).

Social norms can affect compliance because individuals tend to seek the respect of others (McAdams 1997). Social stigma is associated with tax evasion (Scott and Grasmick 1981) and this stigma (and the view of acceptable compliance behavior) varies from country to country (or region to region). Thus, in our model social norms create an additional flow from the evader class to the honest class (beyond the flow created by enforcement). In total, then, evaders feel pressure to become compliant, as a result of both enforcement and social norms. The function $g(x_1)$ represents this combined effect.

In our model, $g(x_1)$ has two components. The first component is β, the proportional flow from the evader class to the honest class arising from enforcement. The second component of $g(x_1)$ is the flow resulting from social norms. On the basis of research in sociology (Hamblin et al. 1973) and following the approach adopted in the literature (e.g., see Lindbeck et al. 1999, 9), we assume that this second component is s-shaped, so that the overall function $g(x_1)$ is also s-shaped as illustrated in Figure 1. The "s" shape of the second component of $g(x_1)$ suggests that when relatively few people adopt a norm, the effect on others is small. As the consensus about a norm increases, its effect on others increases, initially at an increasing rate and later at a decreasing rate. The effect of a social norm on behavior eventually levels off as the proportion of the population adopting the norm approaches 100 percent. The maximum level of social influence is γ.

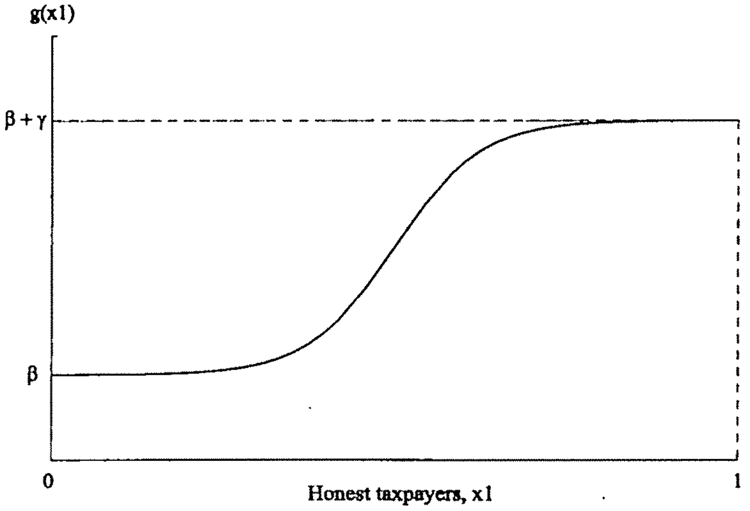
Consequently, our full model incorporates:

- the effect of others' behavior, triggering a flow from honesty to susceptibility;
- the effect of a social norm, causing some evaders to become honest each period;
- an enforcement effect triggering a shift from evasion to honesty; and
- an enforcement effect on the flow from susceptibility to evasion.

---

[11] Social norms influence behavior in a number of settings (e.g., see Grasmick and Green [1980] for work in a legal context, and King [2002] for an experimental study showing that professional norms mitigate auditors' propensity to engage in self-serving bias.)

FIGURE 1
The Proportion of Evading Taxpayers Flowing to the Honest Class, $g(x_1)$



The function $g(x_1)$ represents the total proportion of evading taxpayers flowing to the honest class. The flow consists of two components. A proportion of evaders, β, become honest as a result of enforcement efforts. Additional evaders become honest as a result of social norms. The effect of social norms is s-shaped and depends on the proportion of honest taxpayers, x_1. γ is the maximum proportion of evaders that can flow to the honest class as a result of social norms.

The full model is presented below:

$$\begin{aligned}\dot{x}_1 &= g(x_1)x_3 - rx_1x_3 \\ \dot{x}_2 &= -\alpha x_2 + rx_1x_3 \\ \dot{x}_3 &= \alpha x_2 - g(x_1)x_3\end{aligned} \tag{5}$$

The model in (5) does not rely on traditional game theory, in which the goal is to derive from individuals' rational choices a set of strategy assignments associated with payoffs. The greatest advantage of game theory is that it allows for deduction. The principal shortcomings of a game theory approach are that (1) it cannot explain how the economy changes over time, or how the equilibrium arises, and (2) it focuses on interactions among small numbers of individuals (usually two). In contrast, our dynamic systems analysis stems from an aggregate model that represents the state of the economy by a set of variables, and a system of differential equations describes how these variables change over time. We examine the resulting trajectories. Our approach has much in common with evolutionary game theory focusing on system dynamics (Weibull 1997).[12] However, although our techniques are similar to those used in evolutionary game theory, we do not build on a standard microeconomic foundation. Rather than specifying payoffs, information, and actions for all players, we focus on social-psychological first principles and the consequent evolution of aggregate behavior over time.

---

[12] In accounting, Bloomfield's (1995, 1997) work on the evolution of expectations and its effect on the predictive ability of a Nash equilibrium is in this line of research.

The system of equations in (5) describes the way valid populations evolve over time. In our analysis, we begin by establishing that initially valid populations will remain valid as they evolve. Next, we identify equilibrium points that represent balanced populations that are invariant over time. We then evaluate the stability of these equilibria to determine whether one could reasonably expect that they would ever be reached over time by a set of valid starting populations. We complete our analysis by investigating the effect of changes in enforcement regime on the behavior of the system.

III. AGGREGATE MODEL ANALYSIS

Because (5) is a system of differential equations, their solution is a set of functions of time, $x_1 = x_{1t}$, $x_2 = x_{2t}$, and $x_3 = x_{3t}$ for all values of time t greater than some starting point, $t = t_0$. The aim in analyzing the system is to understand the dynamics of the solutions and how the parameters in the model (e.g., r, α, and β) affect these dynamics. We will show that this particular system of equations has at least one stable equilibrium for each class of taxpayer; i.e., x_{1t} tends to a value, x_1^*, as t increases, regardless of the initial value for x_1. This also holds true for x_2 and x_3.

To begin our analysis, we verify that the solutions are consistent with our interpretation. Since x_1, x_2, and x_3 each represent a partition of the population, and together, these three categories are exhaustive, they must sum to 1 at all values of t. That is:

$$x_1(t) + x_2(t) + x_3(t) = 1. \tag{6}$$

To verify that this condition is true, we begin by assuming that Equation (6) holds true for the initial conditions of our model; i.e.:

$$x_1(0) + x_2(0) + x_3(0) = 1. \tag{7}$$

Then, if we differentiate Equation (6) with respect to t, the result is 0 for all values of t, so that the value of Equation (6) equals a constant with respect to time. Hence, if Equation (7) is satisfied initially, then Equation (7) also will be satisfied for all subsequent values of t.

Next, we establish that, given a set of valid initial values for x_1, x_2, and x_3, the values will remain valid as time increases. Using the system of equations in (5), we develop a set of working equations. Through substitution:

$$\begin{aligned} \dot{x}_1 &= [g(x_1) - rx_1]x_3, \\ \dot{x}_3 &= \alpha(1 - x_1 - x_3) - g(x_1)x_3, \\ x_2 &= 1 - x_1 - x_3. \end{aligned} \tag{8}$$

Before proceeding, we introduce the notion of a *vector field*. Just as a function in one variable associates a value or number with each point on the line, a vector field associates a vector with each point in the plane. One can represent a vector field graphically. If a vector field, φ, is defined on some region of the plane, and if we choose a grid of points that covers the region, then we can represent the vector field with arrows drawn from each point, x, where each arrow represents the vector $\varphi(x)$. The arrow's length equals the norm of $\varphi(x)$ and the arrow points in the direction of the vector. One can view the vector field arising from our system of differential equations as the local set of evolutionary laws.

We can represent the system of differential equations in (5) by a vector field that associates the point (x_1, x_2, x_3) in space with the vector $(\dot{x}_1, \dot{x}_2, \dot{x}_3)$. A solution to the equations $(x_1(t), x_2(t), x_3(t))$ traces out an unbroken curve in space as t increases. It has the property that, at each point (x_1, x_2, x_3) on the curve, the tangent to the curve is the value of the vector field at that point.

Recall that a valid population for our model is a point (x_1, x_2, x_3) in R^2 (two-dimensional Euclidean space) where $x_1, x_2, x_3 \geq 0$ and $x_1 + x_2 + x_3 = 1$. The region V containing all valid populations is bounded by $x_1 = 0$, $x_2 = 0$, and $x_3 = 0$. If all initially valid populations in our system remain valid over time, then the vector field $(\dot{x}_1, \dot{x}_2, \dot{x}_3)$ at any point on the boundary of V is directed toward the interior of V. That is, at each point on the boundary of V:

$$(\dot{x}_1, \dot{x}_2, \dot{x}_3) \cdot n \leq 0, \tag{9}$$

where "·" is the dot product operation and n is the outer normal vector (the vector perpendicular to and outside the valid space, V).

On the plane $x_1 = 0$, the outer normal vector is $n = (-1, 0, 0)$ and the dot product is $-\dot{x}_1 = -g(0)x_3 \leq 0$. Similarly, the outer normal vectors for $x_2 = 0$, $x_3 = 0$, and $x_1 + x_2 + x_3 = 1$ are $n = (0, -1, 0)$, $(0, 0, -1)$, and $(1, 1, 1)$, respectively. The resulting dot products are:

$$-\dot{x}_2 = -rx_1x_3 \leq 0,$$
$$-\dot{x}_3 = -\alpha x_2 \leq 0, \text{ and}$$
$$\dot{x}_1 + \dot{x}_2 + \dot{x}_3 = 0.$$

Because all the dot products are less than or equal to 0, we can conclude that any initially valid population will remain valid as time passes. This result is critical to the plausibility of the model.

Equilibrium Populations

Next we consider the behavior of the system of equations as a function of time. We begin by identifying points where the value of each of the differential equations equals 0, indicating no change over time. These points represent equilibria for the three classes of taxpayers. We find the equilibria by solving:

$$\dot{x}_1 = [g(x_1) - rx_1]x_3 = 0, \text{ and} \tag{10}$$

$$\dot{x}_3 = \alpha(1 - x_1 - x_3) - g(x_1)x_3 = 0 \tag{11}$$

for x_1 and x_3. We can then calculate $x_2 = 1 - x_1 - x_3$. Note that $\dot{x}_2 = -\dot{x}_1 - \dot{x}_3$, so that $\dot{x}_1 = \dot{x}_3 = 0$ implies that $\dot{x}_2 = 0$.

Equation (10) is satisfied either when $x_3 = 0$ or when $g(x_1) = rx_1$. Thus, there are two possible cases to investigate. In the first case, when $x_3 = 0$, Equation (11) becomes $0 = \alpha(1 - x_1)$. Since α is a positive number less than or equal to 1, x_1 must equal 1. Thus, an equilibrium may exist in which all taxpayers belong to the honest class.

In the second case, where $g(x_1) = rx_1$, Equation (11) gives $rx_1x_3 + \alpha x_3 = \alpha(1 - x_1)$, so that:

$$x_3 = \frac{1 - x_1}{1 + \frac{r}{\alpha} x_1}.$$

Substituting the solution for x_3 into the working equation for x_2 in Equation (8) and simplifying yields a value for x_2 of:

$$x_2 = \frac{r x_1 (1 - x_1)}{\alpha + r x_1}.$$

The foregoing analysis leads to the following proposition.

Proposition 1: There are two types of equilibrium populations $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$:
(i) A uniformly honest population $(\hat{x}_1, \hat{x}_2, \hat{x}_3) = (1, 0, 0)$;
(ii) A mixed population of the form:

$$(\hat{x}_1, \hat{x}_2, \hat{x}_3) = \left(\bar{x}, \frac{r\bar{x}(1 - \bar{x})}{\alpha + r\bar{x}}, \frac{1 - \bar{x}}{1 + \frac{r}{\alpha} \bar{x}} \right)$$

where $\bar{x}$ satisfies $g(\bar{x}) = r\bar{x}$.

Location of Equilibria

So far, we have established that equilibrium populations, $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$, may exist and that these populations may be either uniformly honest or mixed. To determine the location of the mixed equilibria, we must identify the values of x_1 where $g(x_1) = r x_1$. In addition, each equilibrium must satisfy the constraints $\hat{x}_1, \hat{x}_2, \hat{x}_3 \geq 0$ and $\hat{x}_1 + \hat{x}_2 + \hat{x}_3 = 1$. Assume that $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ is a type (ii) equilibrium with $0 \leq \hat{x}_1 \leq 1$. Then $1 - \hat{x}_1 \geq 0$. Because α and r are positive, $1 + (r/\alpha)\,\hat{x}_1 \geq 0$. It follows that $\hat{x}_3 \geq 0$ (see the formula for $\hat{x}_3$). Note that:

$$\frac{1 - \hat{x}_1}{1 + \frac{r}{\alpha} \hat{x}_1} \leq \frac{1}{1 + \frac{r}{\alpha} \hat{x}_1} \leq 1. \tag{12}$$

So $\hat{x}_3 \leq 1$. Now, consider the sum:

$$\begin{aligned} \hat{x}_1 + \hat{x}_3 &= \hat{x}_1 + \frac{1 - \hat{x}_1}{1 + \frac{r}{\alpha} \hat{x}_1} \\ &= \frac{1 + \frac{r}{\alpha} \hat{x}_1^2}{1 + \frac{r}{\alpha} \hat{x}_1}. \end{aligned}$$

By hypothesis, each term is not negative, so the sum is not negative. For $\hat{x}_1 \in [0, 1]$, $\hat{x}_1^2 \leq \hat{x}_1$, so:

$$\frac{1 + \frac{r}{\alpha}\hat{x}_1^2}{1 + \frac{r}{\alpha}\hat{x}_1} \leq \frac{1 + \frac{r}{\alpha}\hat{x}_1}{1 + \frac{r}{\alpha}\hat{x}_1} = 1. \tag{13}$$

We can therefore conclude that $\hat{x}_1 \in [0, 1]$ implies that $\hat{x}_1 + \hat{x}_3 \in [0, 1]$. It follows that $\hat{x}_2 \in [0, 1]$. This reduces the search for type (ii) equilibria to finding all $\bar{x}$ satisfying $g(\bar{x}) = r\bar{x}$ and $\bar{x}_1 \in [0, 1]$, leading to the following proposition.

Proposition 2: An equilibrium $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ of type (ii) is valid if and only if $0 \leq \hat{x}_1 \leq 1$ and $g(\hat{x}_1) = r\hat{x}_1$.

The proportion of honest taxpayers is in equilibrium when the proportion of taxpayers flowing into the honest class, $g(\hat{x}_1)x_3$, equals the taxpayers flowing out of the honest class, $r\hat{x}_1x_3$. When the honest proportion is in equilibrium, then both the susceptible and evading classes must also be in equilibrium because the flows among the classes are constant.

Understanding the Dynamics of the System

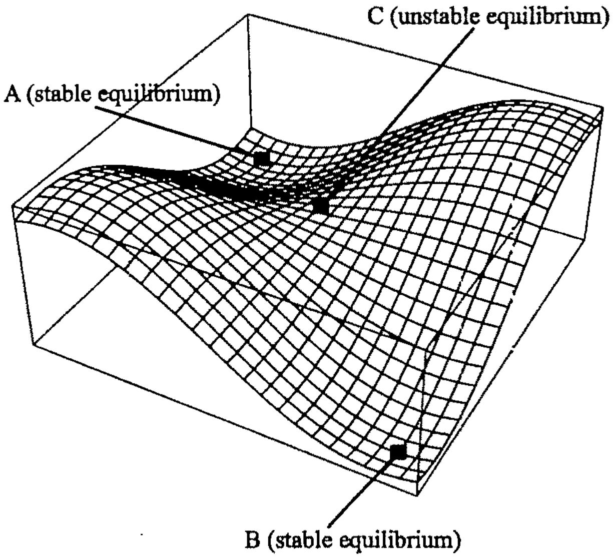
We have established that equilibria exist in the system, but we must address two additional issues to determine whether and under what conditions the equilibria will be reached. First, we must determine whether *basins of attraction* exist (i.e., whether the equilibria attract nearby points). Second, if basins of attraction do exist, we must determine whether they extend to all valid starting populations. If this second condition is satisfied, then all valid populations will eventually reach one of our equilibria, and the dynamics of the system will be fully specified.

An equilibrium point that attracts nearby points over time is *stable*. We define an equilibrium to be stable if there is a neighborhood N of (x^*, y^*) such that if (x_t, y_t) is in the neighborhood N, then $\lim_{t\to\infty} (x_t, y_t) = (x^*, y^*)$ (Arnol'd 1991, 156). For an unstable equilibrium, within even a tiny neighborhood N of the point, some (x_0, y_0) exists in N and some neighborhood U of the equilibrium such that (x_t, y_t) is never permanently in U.

More intuitively, imagine an empty pool with a saddle-shaped bottom, as in Figure 2. If one were to drop a ball in the pool, it would likely roll toward one of the deep ends, eventually coming to rest at one of the two deepest points (labeled A and B, respectively). These points at which the ball stops moving represent stable equilibria. They are stable in that, if the ball is in the point's "neighborhood," (i.e., the sloped area that surrounds the point), then the ball will be attracted to that particular point. Alternatively, the ball could be carefully balanced at point C, the lowest point in the shallow section (i.e., not on either slope to the deep ends), and remain at rest. This point is also an equilibrium; however, it is unstable. A small movement of the ball away from this point would cause it to roll to one of the deep ends, eventually coming to rest at A or B.

We investigate the stability properties of our system of differential equations to gain insight into the behavior of the system over time. For example, given a stable equilibrium and selected initial conditions close enough to this point, the solution path to a system of differential equations will tend toward and eventually reach equilibrium. Alternatively, a selected initial condition may lie near an unstable equilibrium. In this case, the system may

FIGURE 2
Basins of Attraction and Stable vs. Unstable Equilibria



This surface illustrates stable and unstable equilibria. A ball placed on the surface near point A (or B) will roll toward the point and eventually come to rest there. Thus, A and B are stable equilibria. A ball placed near C may initially roll toward C, but it will not come to rest there (although a ball carefully balanced at C will remain at rest). C is an unstable equilibrium.

initially tend toward, but not stabilize at, the critical point. Consequently, there would be no reason to expect the equilibrium to obtain.

To ascertain stability in our system, we evaluate a linear approximation of the system near an equilibrium. For the case of an honest equilibrium, our analysis (formally presented in the Appendix) leads to the following proposition regarding stability.

Proposition 3: Assume that $g(x)$ has a convergent Taylor series expansion near $x = 1$ and $g(1) > 0$. The equilibrium $(1, 0, 0)$ is stable if and only if $r < g(1)$.

Proof. See the Appendix.

Proposition 3 is intuitive in the neighborhood of an honest equilibrium, where $g(1)$ captures the effect of social norms and enforcement on the conversion of evaders into honest taxpayers. When the flow of honest taxpayers into the susceptible class (via the infection rate, r) is less than the flow from the evading class into the honest class (via a value near $g(1)$), then all taxpayers will eventually become honest.

An analysis of the stability of the mixed equilibrium leads to the following proposition.

Proposition 4: Assume that for any $\bar{x} \in [0, 1]$, $g(x)$ has a convergent Taylor series near $\bar{x}$ and $g(\bar{x}) \geq 0$. Let $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ be a mixed equilibrium. The equilibrium is stable if and only if $r > g'(\hat{x}_1)$.

Proof. See the Appendix.

To understand the intuition underlying Proposition 4, recall that, by Proposition 1, the proportion x of honest taxpayers is in equilibrium when the proportion rx of infected taxpayers reduces the honest population at the same rate that social forces and enforcement (g(x)) increase it. This equilibrium will be stable if proportions below x are driven upward and proportions above x are driven downward. This happens when $g(x) > rx$ for points below x and when $g(x) < rx$ for points above x. These inequalities hold only if $g'(x) < r$.[13]

Determining the Domain of Stability

Having evaluated the conditions under which equilibria are stable, we turn our attention to the second issue: whether and under what conditions the domain of stability encompasses the entire set of valid populations (global stability). In evaluating nonlinear dynamic systems, one can rarely derive equations for the domain of stability (Lyapunov functions), and unfortunately this holds true for our case as well. However, using a combination of graphics, computational analysis, and an appeal to the Poincaré-Bendixson theorem (P-B theorem), we provide evidence that all valid populations will eventually reach one of the stable equilibrium points.

The P-B theorem applies when the vector field is planar and bounded, and equilibria are isolated. When these conditions are met (as in our system), either (1) the limit set (the range of points possible in a neighborhood as a result of the system dynamics) or a subset of the limit set is an equilibrium, or (2) the limit set is a periodic orbit. To ascertain which of these outcomes applies to our system, we turn to a qualitative evaluation of the system.

Recall from Proposition 1 that honest equilibria always exist, and from Proposition 2 that mixed equilibria exist only where x_1 is between 0 and 1 and the line rx_1 and the function $g(x_1)$ intersect. In Panel A of Figure 3, we provide an example in which three equilibria exist: two mixed equilibria at points a and b and an honest equilibrium at point c.

We can evaluate the stability of these equilibria using the results of Propositions 3 and 4. According to Proposition 3, a completely honest equilibrium is stable if and only if $r < g(1)$. That is, the honest equilibrium is stable when the flow of honest taxpayers into the susceptible class is less than the flow of evaders into the honest class. This condition is satisfied in Panel A of Figure 3, so the honest equilibrium c is stable.

For a mixed equilibrium to be stable, populations below the equilibrium must be driven upward toward the equilibrium and populations above the equilibrium must be driven downward toward the equilibrium. This happens when $g(x) > rx$ for points below x and when $g(x) < rx$ for points above x. According to Proposition 4, these inequalities hold only when the slope of the line rx_1 exceeds the slope of $g(x_1)$ at the intersection point (i.e., when the proportions below the equilibrium are driven upward and proportions above the equilibrium are driven downward). This condition is met at a, but not at b. Thus, a is a stable equilibrium value for the honest population proportion (x_1) and b is unstable.

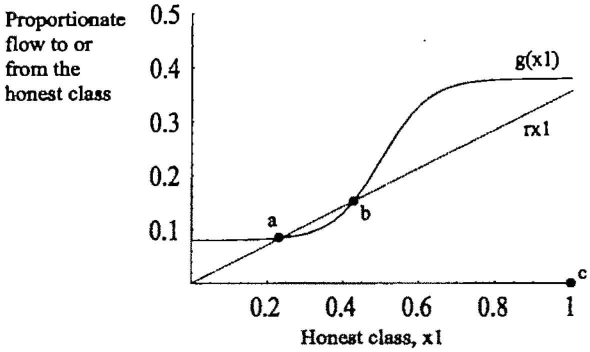
Having identified the honest-population proportions (x_1) in our three equilibria in the example, we can use the results of Proposition 1 to compute the proportion of evaders (x_3) and susceptible taxpayers (x_2) for each mixed equilibrium. We graphically represent these equilibria in the (x_1, x_3) plane in Panel B of Figure 3. The black dots represent stable equilibria (at a and c) and the gray triangle (at b) represents an unstable equilibrium.

---

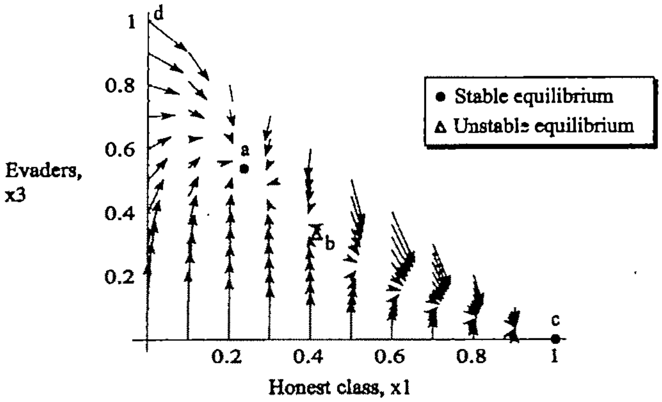
[13] We thank one of our anonymous reviewers for this insight.

FIGURE 3
The Location of Equilibria, Their Stability, and the Movement of Out-of-Equilibrium Populations over Time

Panel A: Location of honest equilibrium values.



Panel B: Stable and unstable equilibria in the vector field.



Panel A illustrates how we identify equilibrium values. Mixed equilibria exist when the number of evading taxpayers moving to the honest class equals the number of honest taxpayers moving to the susceptible class (when $g(x_1)$ equals rx_1, at points a and b). An honest equilibrium always exists. The honest equilibrium is stable when the flow out of the honest class is less than the flow into the honest class (when the value of r is less than the value of g(1)). In Panel A, this condition is satisfied for point c. Mixed equilibria are stable if the slope of rx_1 exceeds the slope of $g(x_1)$ at the intersection point. This condition is met at a, but not at b.

Panel B shows the equilibrium values on the (x_1, x_3) plane, together with the associated vector field. The arrows in the vector field illustrate the movement of points over time toward the stable equilibria, a and c. This figure illustrates only one of several possible relations between rx_1 and $g(x_1)$.

Panel B of Figure 3 also uses arrows to illustrate the vector field associated with the system. The vector field here is a two-dimensional representation of a saddle-shaped surface similar to that depicted in Figure 2. We constructed the vector field computationally by determining the dynamic solution to the system for numerous starting populations and representing them graphically as arrows.[14] The direction of each arrow shows the path the system takes over time from the arrow's origin. The length of the arrow represents the slope of the surface (longer arrows are equivalent to a steeper slope). To illustrate, given an initial population in which everyone evades (at d in Panel B of Figure 3), the population would move to point a over time and then remain there (in equilibrium). We can observe the dynamics by following the path suggested by the arrows, beginning at d and ending at a. From the figure (and the P-B theorem), one can see that valid populations (except for the unstable equilibrium at b) appear to move toward either a or c. Hence, our analysis suggests that the limit set in the system is either a, b, or c (i.e., the system appears to exhibit global stability). We observe no evidence of a periodic orbit in the vector field.

Figure 3 depicts only one possible relation between the line rx_1 and the function $g(x_1)$. We evaluated 11 possible relations that arise from combinations of differently sloped lines rx_1 and differently shaped $g(x_1)$. These 11 relations are exhaustive in that they describe all types of possible equilibria and the related dynamics that could occur in our system. In every case, either one or two stable equilibria exist and we observe no evidence of a periodic orbit.

The Effect of Enforcement on Evasion

Enforcement affects the movement of susceptible taxpayers to the evading class (α) and the movement of evaders to the honest class (β). To better understand the effect of enforcement in our model, we begin by examining benchmark cases for these two flows. For the case in which $\alpha = 0$, susceptible taxpayers remain susceptible and never become evaders (presumably because of draconian enforcement), by the system of differential equations in (5). As long as at least some evaders are initially present in the system, there is a flow out of the evading population and into the honest population, until, in the limit, all taxpayers comply with the law (being either honest or susceptible). This result obtains even if there are no honest taxpayers initially, because $g(0) = \beta > 0$. In equilibrium, as α decreases (and perceived enforcement severity increases), by Proposition 1, the proportion of evaders decreases at an increasing rate and the proportion of susceptible taxpayers increases.

An analysis of β is more complex. In an initial condition in which there are no honest taxpayers ($x_1 = 0$) and no enforcement efforts ($\beta = 0$), $g(x_1)$ has no effect. There is no flow from the evading class to the honest class. In the limit, the entire population evades (assuming $\alpha > 0$). However, if some proportion of the initial population is honest ($x_1 > 0$), then some evaders become honest in each period, as a result of social norms. Because β affects only the intercept and not the shape of $g(x_1)$, the system dynamics and the limit set depend on the relation between r and $g(x_1)$, as noted previously.

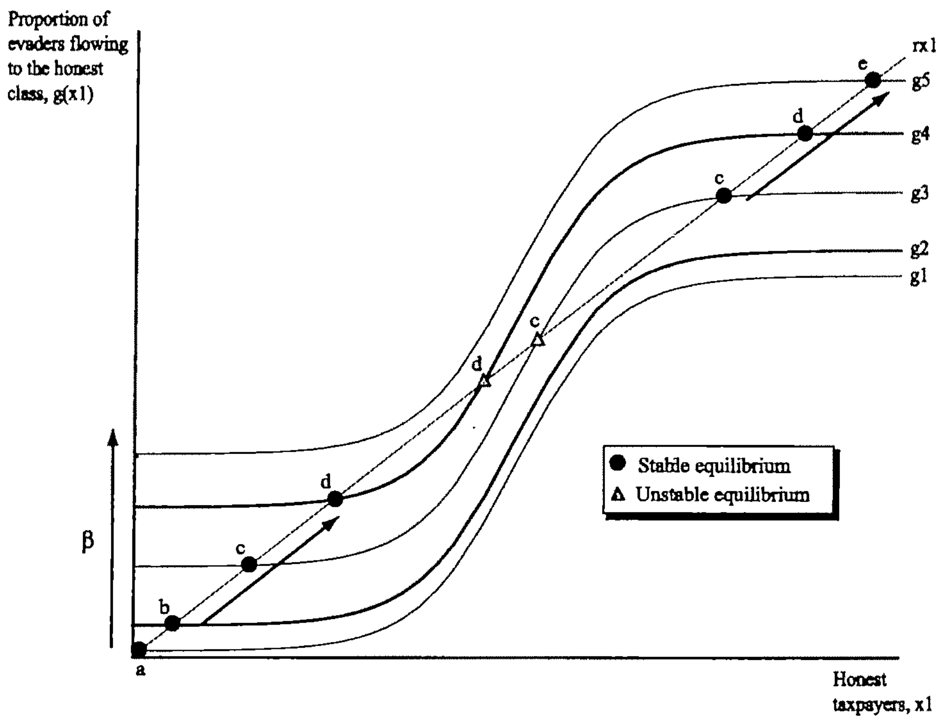
Figure 4 illustrates the effect of increasing the enforcement-related flow from the evading class to the honest class, β, on the equilibrium values for x_1. At low levels of β (on curves g1 and g2), the proportion of honest taxpayers is low (at stable equilibrium points a and b). As β increases, the curve moves from g2 to g3. At some point between these curves, $g(x_1)$ will be tangent to rx_1. Beyond this point, the line and the curve will intersect in three places. These intersection points, c, are identified on g3. The high and the low points are stable equilibria and the middle point is an unstable equilibrium. The unstable

---

[14] The qualitative aspects of the vector field obtain for any $g(x)$ satisfying the relations in Panel A. For computation purposes, we use the function $g(x) = \beta + (4a/d(1 + e^{db}))[(e^{db} - e^{-d(x-b)})/(1 + e^{-d(x-b)})]$ where the equality $\gamma/a = 4(1 - e^{-d})/(d(1 + e^{-db})(1 + e^{-d(1-b)}))$ is satisfied if and only if $a > \gamma > 0$. A proof showing that this equation satisfies the criteria required for $g(x)$ is available from the authors.

FIGURE 4

The Effect of Increasing the Enforcement-Driven Flow from the Evading Class to the Honest Class, β, on the Equilibrium Proportion of Honest Taxpayers, x_1



As the enforcement-driven flow from the evading class to the honest class, (β), increases, the total proportion of evading taxpayers moving to the honest class, $g(x_1)$, moves upward from curve gl to g5. When β is small (gl), only one stable equilibrium exists at a. In this equilibrium, only a few taxpayers are honest. Increasing enforcement to g2, the proportion of honest taxpayers in equilibrium increases to b. As enforcement continues to increase (g3), three equilibria exist, all denoted by c. The high and low intersection points are stable equilibria and the middle point is unstable. At a higher level of enforcement (g4), the proportion of honest taxpayers increases under both stable equilibria but decreases in the unstable equilibrium (all denoted by d). At high levels of enforcement (g5), only one stable equilibrium exists, denoted by e. Most taxpayers in this equilibrium are honest.

equilibrium plays an important role in determining the dynamics of the system. The vector field in Panel B of Figure 3 indicates that the system is characterized by a saddle-shaped surface. The unstable equilibrium is on the saddle, so that out-of-equilibrium populations above the unstable equilibrium will move to the high equilibrium point, while out-of-equilibrium populations below the unstable equilibrium will move to the low equilibrium point over time. As β continues to increase and $g(x_1)$ moves to g4, the proportion of honest taxpayers increases in the two stable equilibria and decreases in the unstable equilibrium (at d). As the unstable equilibrium decreases, out-of-equilibrium populations that would have moved to the lower stable equilibrium under g3 begin, under the greater enforcement depicted in g4, to move to the upper stable equilibrium instead. Finally, as $g(x_1)$ moves from g4 toward g5, the line and curve will again reach a point where they are tangent. Beyond this point, they intersect only once. This intersection occurs on g5 at e. In this (stable) equilibrium, most taxpayers are honest. This figure also highlights the importance of the s-shaped function describing the flow of taxpayers from the evading class to the honest class; multiple equilibria arise as a result of the nonlinear effect of social norms at some levels of β.

Figure 5 summarizes our analysis.[15] As enforcement increases, β increases. The solid curves show how an increase in β affects the evading proportion in equilibrium. Based on our analysis of Figure 4 (viz., curves g1 and g2), at low levels of enforcement (below A in Figure 5), the proportion of honest taxpayers is small in equilibrium. By Proposition 1, x_1 and x_3 are negatively correlated, so the proportion of evaders is relatively high in these equilibria. A and B in Figure 5 correspond to the values of β where rx_1 and $g(x_1)$ are tangent (between curves g2 and g3 and again between curves g4 and g5 in Figure 4). At intermediate levels of enforcement (between A and B), two stable equilibria exist, admitting the possibility of either a mostly evading population or a mostly compliant population. This result follows from curves g3 and g4 in Figure 4. At sufficiently high levels of enforcement (above B), the system returns to a single, mostly honest equilibrium.[16] Curve g5 in Figure 4 illustrates this result.

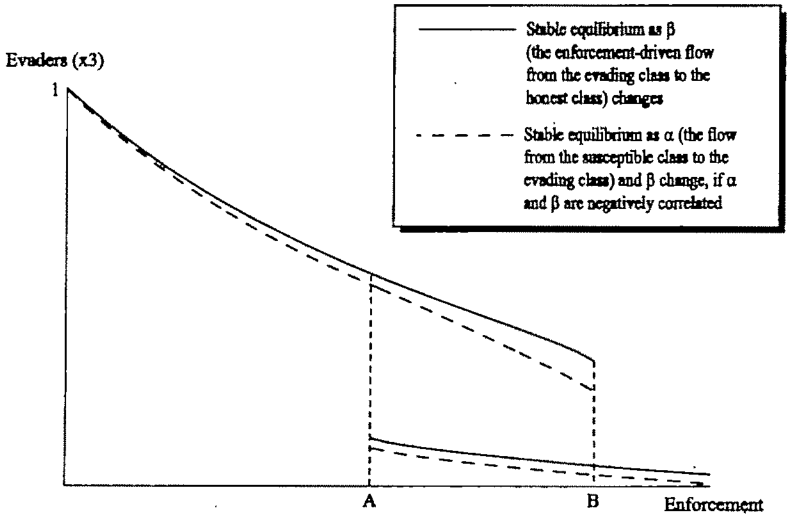
---

[15] We constructed Figure 5 from a numerical analysis, using the equation in footnote 14 as g(x).

[16] For some parameterizations of rx_1 and $g(x_1)$, the lower curve in Figure 5 reaches 0, or a completely honest population.

FIGURE 5

The Effect of Enforcement on Taxpayer Evasion



The solid curves in the figure show stable equilibrium values as enforcement changes, assuming that enforcement affects only β (the flow from the evading to honest class). The dashed curves show stable equilibrium values if enforcement affects both β and α (the flow from the susceptible to the evading class). Below A and above B, there is a unique equilibrium value for each level of enforcement. Between A and B, two stable equilibria exist. If a population in equilibrium at point A on the lower curve experiences a decrease in enforcement, then it will move to a new equilibrium along the upper curve. Similarly, populations at point B on the upper curve will move to a new equilibrium on the lower curve if the tax agency increases enforcement.

An increase in enforcement also decreases the proportion of susceptible taxpayers in our model who become evaders (α). By Proposition 1, when α decreases, the number of evaders decreases at an increasing rate. This relation holds for all stable, mixed equilibria. Thus, α reinforces β, further increasing the responsiveness of the population to enforcement efforts. The dashed curves in Figure 5 illustrate the joint effect of enforcement through both α (the reduced flow from the susceptible to the evading class) and β (the increased flow from the evading to the honest class).

We now describe the dynamic behavior of initially out-of-equilibrium populations. Because our analysis suggests global stability, at low and high levels of enforcement, all populations will move toward and eventually reach the unique equilibrium value on the curve. As noted previously, the unstable equilibrium value determines which equilibrium a population will reach. Populations with honest proportions smaller than the unstable equilibrium value will move to the noncompliant equilibrium, while populations with honest proportions greater than the unstable equilibrium will move to the compliant equilibrium.

We can make several interesting observations using Figure 5. Consider a population that has reached a unique equilibrium value on the upper curve, where enforcement is below A and evasion is relatively widespread. If we repeatedly increase enforcement in very small increments and allow the population to adjust to a new equilibrium after each change, it will follow the path from left to right suggested by the upper curve to point B. These increases in enforcement reduce evasion (evidenced by the downward slope of the curve), but a large proportion of the population will still evade. When enforcement exceeds B by an infinitesimal amount, the population will move to the lower curve in equilibrium, resulting in a mostly compliant regime. Thus, a small increase in enforcement leads to a dramatic shift from aggregate noncompliance to aggregate compliance. From a policy perspective, this result suggests that, in regions or time periods where evasion is widespread, increased enforcement may initially have a limited effect on compliance. At some point (B in our figure), however, enforcement can trigger a more profound change in the population's behavior, leading to a mostly compliant regime. Once this occurs (and the population reaches an equilibrium on the lower curve), the tax agency can slowly relax enforcement efforts (down to A in the figure) with little change in the number of evaders. The population will follow the equilibrium suggested by the lower curve if we allow it to adjust to a new equilibrium after each small change. However, an infinitesimal reduction in enforcement below A will again trigger a dramatic change. The population will move to a new equilibrium on the upper curve, where many more taxpayers evade. Attempts to remedy the resulting evasion "epidemic" by increasing enforcement to a point between A and B, which was previously sufficient to maintain compliance, would be largely unsuccessful because the population would again follow the upper curve until enforcement reaches B.

IV. THE AGENT-BASED MODEL

Our aggregate-level model has several shortcomings. First, to maintain tractability, we must make unrealistic assumptions regarding movement among classes of taxpayers. For example, we assume that the flows between classes are in constant proportions (even though there are likely to be random fluctuations in the real world, introduced by randomness in audit frequency and individual behavior). Second, we assume that social norms have an s-shaped effect on the movement from the evading class to the honest class. Third, the aggregate-level model does not explicitly account for micro-level taxpayer behavior. To address these limitations, we supplement our analysis with a micro-level agent-based computational model.

Background

Agent-based modeling has roots in biology (e.g., Crist and Haefner 1994; Haefner and Crist 1994) and is becoming a more popular method to address social sciences issues[17] as advances in computing technology have made these models feasible.[18] Agent-based models explicitly characterize individual agents as collections of internal states and endowments (e.g., knowledge and assets). Some agent states and endowments are fixed for the life of the agent, whereas others change as a result of interactions with other agents or the environment. In addition to agents, agent-based models also characterize the environment in which agents interact. An environment might represent a landscape or lattice of resource-bearing sites in ecological or anthropological research, a set of institutional (market) rules in economics, or a communication network in a society (in sociology). Agents can either interact with their environments (e.g., travel through their environments to collect resources) or with each other through their environments (e.g., engage in trade) via a set of behavioral and institutional rules.

After defining agents, their behaviors and endowments, and the environment, we allow the model to evolve. Agent-based models focus on aggregate-level behaviors that emerge endogenously from agent interaction. In particular, researchers use agent-based models to identify micro-level characterizations that are sufficient to generate macroscopic social structures or collective (aggregate) behaviors. Put another way, agent-based models grow collective behaviors "from the bottom up" (Epstein and Axtell 1996).

Agent-based models have both advantages and disadvantages (see Axtell [2000] for a more detailed discussion). On the positive side, it is easy to introduce heterogeneity and avoid an appeal to a representative agent, and it is easy to limit agent rationality. In addition, in agent-based models, there is no need to focus on equilibria because the entire dynamic history of the process is available to the researcher. Finally, agent-based models allow investigation of the effects of physical location and social networks.

The most significant disadvantage of agent-based models relative to traditional mathematical models relates to robustness of results. Although each run of an agent-based model amounts to a sufficiency theorem, the model provides no information as to whether the set of assumptions is necessary for the observed results. As Axtell (2000, 6) noted, agent-based models can provide only limited insight into the question: "Given that agent model A yields result R, how much change in A is necessary in order for R to no longer obtain?" Analytical models can usually formally answer this question, but agent-based models can only address it via multiple runs, with systematic variation in initial conditions and parameters. Even then, there are limits to the number of model variations that we can examine (although the continued increase in computing power reduces this limitation).

Creating a Society of Taxpayers

Our agent-based model is based on the framework developed by Gaylord and Davis (1999), which provides a general approach for creating nonspatial agent-based models using

---

[17] Schelling (1969, 1971, 1978) pioneered the use of agent-based models in social science research, using the method to study the formation of segregated neighborhoods. More recently, agent-based models have provided new insights in economics (e.g., see Marimon et al. 1990; Albin and Foley 1992; Marengo and Tordjman 1996; Arthur et al. 1997; Weisbuch et al. 1998), sociology and social psychology (Carley 1991; Glance and Huberman 1993; Danielson 1996; Page and Hong 1996), anthropology (e.g., Kohler et al. 1999), and political science (e.g., Axelrod 1997; Kollman et al. 1997).

[18] For example, the small experiment reported in this paper took approximately 720 hours of processing time on a 550 MHz personal computer.

Mathematica. We model a society of taxpayers, represented by a matrix (in which each row represents the attributes of a single taxpayer) and a set of transformation rules (representing behavioral rules and the environment through which taxpayers interact). Our model evolves by iteratively mapping the transformation rules to the matrix.[19]

Each taxpayer in our model has several attributes, beginning with a unique name. In addition, since taxpayers update their beliefs about enforcement severity and social norms based on the experiences and behaviors of their acquaintances, we specify a randomly determined set of acquaintances for each taxpayer. The model assumes that acquaintance relationships are symmetric (i.e., if taxpayer 1 knows taxpayer 2, then taxpayer 2 also knows taxpayer 1). Each taxpayer's list of acquaintances is fixed.

In the initial condition for our society, we randomly assign one of two behaviors (honest or evading) to each taxpayer. As a society evolves, the taxpayer's behavior may change as a function of his characteristics and his acquaintances' beliefs and behaviors. Honest taxpayers may become susceptible to evasion by viewing the reporting behavior of a randomly chosen acquaintance. If the randomly chosen acquaintance is an evader, then the taxpayer becomes susceptible with some probability (randomly and independently determined for each taxpayer). This probability represents the agent's susceptibility to "infection," analogous to r in the aggregate model, except that the level of susceptibility is a taxpayer characteristic rather than an aggregate flow.

In our model, the tax agency can imperfectly distinguish evaders from compliant taxpayers, so the agency audits evaders with a higher probability. Taxpayers form a belief regarding the severity of enforcement by observing the frequency with which their acquaintances are audited.[20] Susceptible taxpayers will choose to evade if their perception regarding enforcement severity is below a randomly determined threshold[21] (determined independently for each taxpayer). This agent-level threshold captures the effect of the taxpayer's risk preferences and the expected benefits of evasion, which we assume are heterogeneous over our population.

In addition to forming a belief about the severity of the enforcement regime, taxpayers in the agent-based model also form a belief regarding the extent to which honesty represents a social norm. We implement this belief as the proportion of honest taxpayers in each agent's list of acquaintances. As in our aggregate model, the social norm can cause an evader to become honest. In our agent-based model, evading taxpayers convert to honest behavior if (1) they are audited, (2) their perception of enforcement severity equals or exceeds their threshold, or (3) their perception regarding honesty as a social norm exceeds a randomly determined threshold (determined independently for each taxpayer). Table 1 summarizes the agent-based model.

We designed the agent-based model to relax our earlier simplifying assumptions regarding the movement between classes of taxpayers. Consider, for example, our aggregate model's assumption that flows between classes are uniform and proportional. In each time step of our agent-based model, even if every honest taxpayer knows an evader, movement

---

[19] A detailed description of the *Mathematica* code for our model is available from the authors.

[20] Although the base rate of tax audits is common knowledge in some countries (e.g., the U.S. press reports overall audit rates), individuals do not respond normatively to such information. For example, they tend to underweight base rates (Kahneman and Tversky 1973). In addition, other factors affect individuals' likelihood judgments, including the ease with which they can recall occurrences of the event (Tversky and Kahneman 1973). Our assumption that taxpayers form their beliefs on the basis of their own experience and on the basis of their acquaintances' experiences is consistent with people's tendency to underweight base rates and base judgments of likelihood on ease of recall.

[21] For simplicity, we draw all randomly determined values from a uniform distribution.

TABLE 1
Features of Our Agent-Based Model: Taxpayer Attributes and Schedule of Taxpayer Actions

Panel A: Taxpayer Attributes

1. Unique identification number	Each taxpayer's identification number, used as a reference when accessing acquaintance lists.
2. Acquaintance list	A randomly determined, unchanging list of other taxpayers that a given taxpayer consults when updating beliefs regarding enforcement and social norms.
3. Current reporting behavior	The taxpayer's reporting behavior in the current period (honest, susceptible, or evading).
4. Perception of enforcement	The mean audit rate among the taxpayer's acquaintances in the previous period.
5. Perceived social norm	The proportion of the taxpayer's acquaintances who were honest in the previous period.
6. Enforcement threshold	A randomly determined number between 0 and 1 (determined independently for each taxpayer at time step 0). A susceptible taxpayer will evade if her perception regarding the severity of enforcement (attribute 4) is less than this number.
7. Susceptibility parameter	The probability with which a taxpayer will become susceptible if he observes an acquaintance evade. This parameter is a random number between 0 and 1, determined independently for each taxpayer at time step 0.
8. Audit indicator	Indicates whether the taxpayer was audited in the previous time step.
9. Norm threshold	A randomly determined number between 0 and 1 (determined independently for each taxpayer at time step 0). An evading taxpayer will become honest if her perception regarding social norms (attribute 5) is more than this number.

Panel B: Schedule of Taxpayer Actions in Each Time Step

1. Each taxpayer begins a time step by updating her reporting behavior on the basis of a set of rules: (a) honest taxpayers become susceptible with some probability if a randomly selected acquaintance evades; (b) susceptible taxpayers become evaders if perceived enforcement severity is less than their thresholds; (c) evaders become honest if perceived enforcement severity exceeds their thresholds, or if the perceived social norm exceeds their thresholds.
2. The tax agency randomly selects taxpayers to audit, where the probability of auditing an evader is higher than the probability of auditing others. The audit indicator for each taxpayer (attribute 8) is updated to reflect this action. Taxpayers who are not honest become honest upon audit.
3. Each taxpayer updates her perception of social norms (attribute 5) by polling the reporting behavior of taxpayers in her acquaintance list.
4. Each taxpayer updates her perception of the severity of the enforcement regime (attribute 4) by ascertaining the audit frequency of taxpayers in her acquaintance list.

to the susceptible class is *randomly* determined on the basis of each agent's susceptibility. Therefore, random proportions of honest taxpayers become susceptible, and it is possible that all agents could either become susceptible or remain honest. Similar randomness arises in other flows between taxpayer classes. Second, by modeling the decisions of individual agents, we avoid the assumption that social norms have an s-shaped effect on movement from the evading class to the honest class. That is, we describe the model at the level of individual agents, with no attempt to mathematically aggregate individual agent behaviors. Consequently, there is no guarantee that our earlier predictions that hinged on the s-shape of g(x) (i.e., on the nonlinear effect of social norms) will hold in our agent-based model.

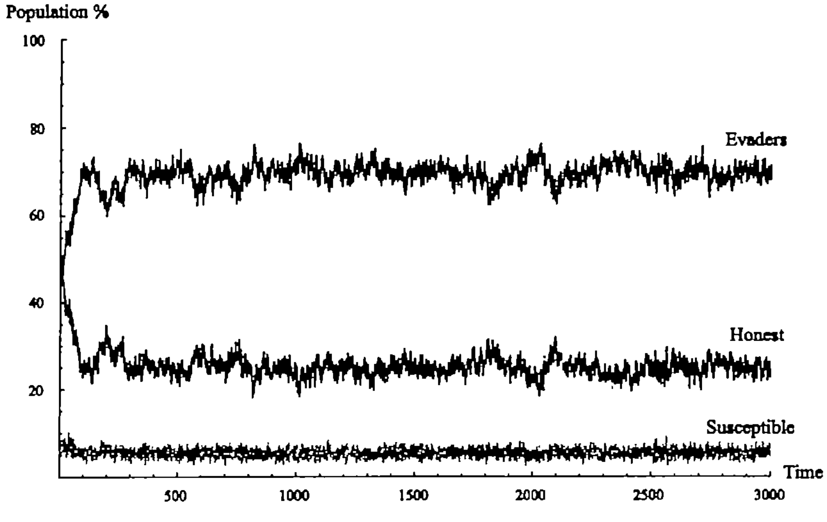
Agent-Based Analysis and Results

In light of our aggregate model results, two properties of the agent-based model are of particular interest. First, do stable equilibria exist? Second, do the comparative static properties of our agent-based model correspond with the predictions in the aggregate model?

To solve the agent-based model, we iterate it forward through time, observing the aggregate dynamics. A typical time series illustrating changes in reporting behavior appears in Figure 6. The population rapidly moves from its initial condition (50 percent evading) toward a value of about 70 percent evading. Then, for the remaining periods, the proportion of the population evading fluctuates around 70 percent. Although this behavior does not precisely match our stability prediction, it is consistent with an alternative notion of stability in which an equilibrium is considered stable if points in the neighborhood of the equilibrium

FIGURE 6

Agent-Based Model: Typical Time-Series Evolution (over 3,000 time steps) of a Taxpayer Population That Was Initially 50 Percent Compliant



Taxpayer compliance behavior aggregated from our agent-based model does not reach a constant state as predicted in the aggregate model. Instead, compliance behavior varies randomly around specific values (roughly 70 percent evaders, 25 percent honest, and 5 percent susceptible). This randomness arises because the agent-based model relaxes the assumption of constant flows among classes of taxpayers.

tend to remain in the neighborhood over time (Lyapunov stability). Thus, our earlier inference that equilibria tend to attract nearby points appears to generalize to the agent-based model.[22]

To examine whether the comparative static properties of our aggregate model generalize to the agent-based model, we created nine matched pairs of taxpayer societies. The members of each pair were identical, except for the starting proportion of tax evaders, which we manipulated within pairs at 10 percent and 50 percent. In addition, we manipulated the evader audit likelihood from 0.002 to 0.030 in 0.002 increments for each of the 18 societies.[23] Each society had 500 taxpayers and each taxpayer initially had five acquaintances (after we applied symmetry, the average number of acquaintances was approximately nine). We allowed each society to evolve at each audit rate for 2,000 time steps.

Figure 7 depicts the proportion of evaders in the final time step for each pair of starting populations, by audit rate. For the most part, the graphs resemble Figure 5, corroborating our earlier aggregate model analysis. A sudden increase in compliance is evident in noncompliant populations as enforcement increases (between 20 percent and 60 percent of the population changes from evading to complying in response to a 0.002 increase in audit rates). Likewise, compliant populations are compliant unless enforcement is sufficiently lax, at which point there is a sudden transition to noncompliance. This happens in all but one of the graphs in Figure 7, with increases in evasion ranging from 35 percent to more than 70 percent in response to a 0.002 decrease in audit rates. The figure highlights these sudden changes in equilibrium compliance with darkened segments on the audit-rate axis.[24] The results from our agent-based model differ from Figure 5 in one important way. Rather than observing a "mostly honest" population at sufficiently high audit rates, we observe total honesty. Our aggregate model predicts this result if the rate at which taxpayers become susceptible, r, is less than g(1) when the enforcement-related flow from the evading class to the honest class, β, equals 0.

In summary, our agent-based model supports the predictions from our aggregate-level analysis. We find evidence supporting the existence of multiple equilibria at some enforcement levels (conditioned on the compliance level of starting populations). In addition, the effect of changing enforcement severity in our agent-based model is consistent with predictions from the aggregate model. At critical points, small changes in enforcement can lead to large changes in compliance.

V. CONCLUSIONS

In this paper we examined the effects of others' behavior, enforcement, and social norms on the dynamics of taxpayer compliance. We employed two novel modeling methods. In our aggregate model, we found that two stable equilibria occur at some enforcement levels. The effect of changing enforcement levels on taxpayers' compliance depended on whether the initial population was relatively compliant or noncompliant: (1) in a population that is initially compliant, the tax agency can reduce enforcement down to a critical point with little effect on compliance, and (2) in a population that is initially noncompliant, increasing enforcement leads to modest increases in compliance up to a critical point. In the case of

---

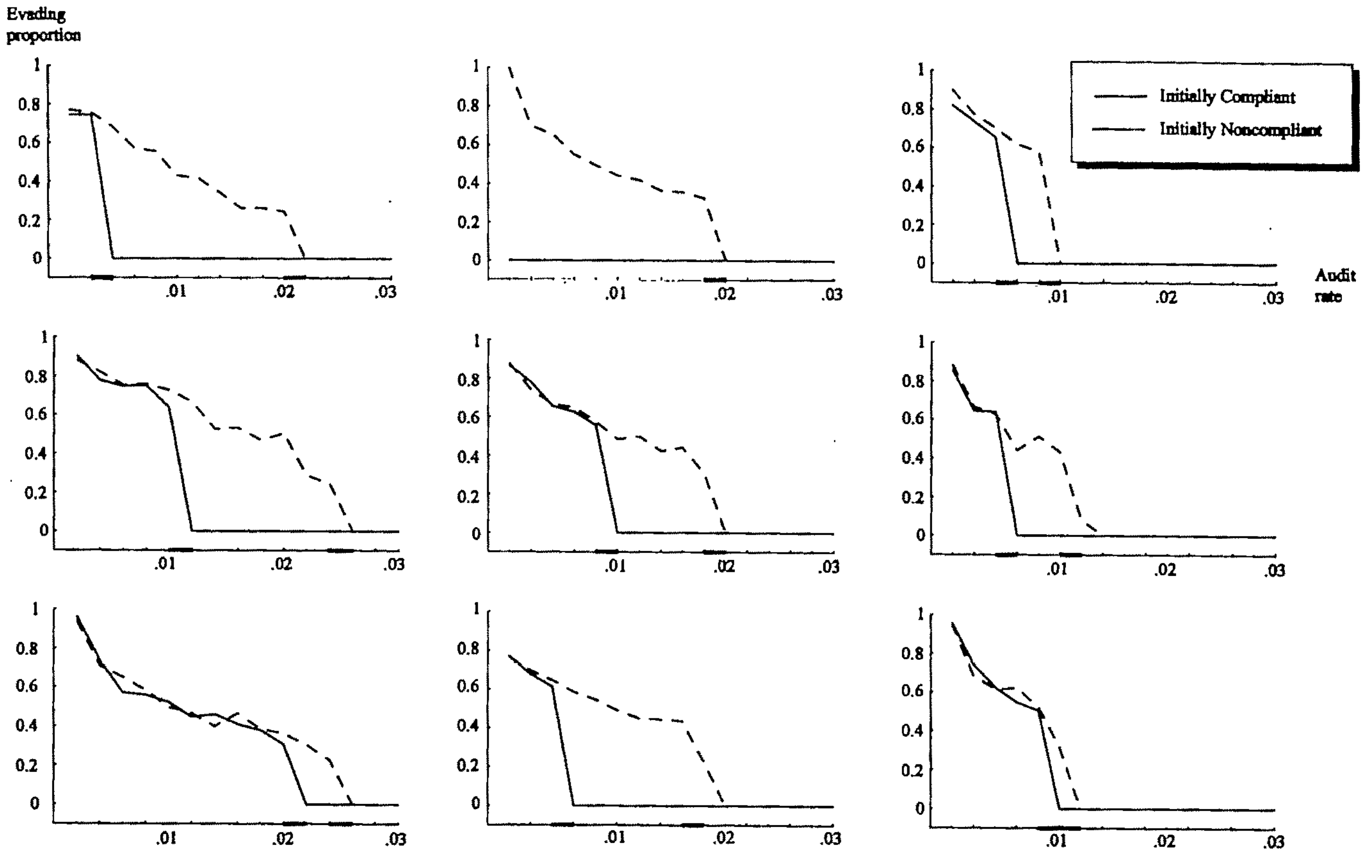
[22] Because agent-based analysis is data driven, it can only falsify (and not verify) predictions. Although the data arising from our analysis are consistent with a Lyapunov-stable system, we cannot prove that our model is Lyapunov stable.

[23] The tax agency did not audit compliant taxpayers in this series of simulations.

[24] We investigated finer partitions (at 0.001 intervals) of audit rates around the sudden transition in behavior, and our results remain consistent with the aggregate model.

FIGURE 7

Evasion Across Audit Rates When Initial Populations Are Either Compliant or Noncompliant in the Agent-Based Model



Plot of the level of taxpayer evasion by audit rate after 2,000 time steps, for nine pairs of starting populations, where the members of each pair differ only in the level of compliance in the initial condition (10 percent evading vs. 50 percent evading). The darkened segments on the audit-rate axis indicate intervals where we observed sudden changes in equilibrium values.

a compliant equilibrium, however, our analysis indicated that when the tax agency decreases enforcement to just below some critical point, a dramatic shift in behavior will occur. The population will undergo a sudden and profound change, moving from compliance to noncompliance in equilibrium. Once this change occurs, attempts to remedy the sudden proliferation of evasion by increasing enforcement to the level in place before the change in behavior will be unsuccessful; the population will remain noncompliant. Similarly, if the tax agency increases enforcement above some critical point in an initially noncompliant population, the population will suddenly become compliant in equilibrium. Then, reducing enforcement to its previous level, below the critical point, will not lead to a significant increase in evasion. These results were robust; evaluation of more complex models with additional flows between taxpayer classes yielded equivalent results.

In contrast to our aggregate model, the agent-based model characterized individual decisions rather than aggregate behavior. Despite relaxing assumptions regarding movement between taxpayer classes and the effect of social norms, the agent-based model's results are generally consistent with the predictions of the aggregate model.

The models yield two major insights. First, in mostly honest populations, the tax agency might view enforcement as a mechanism to prevent evasion epidemics rather than as a way to improve existing compliance. Second, the possibility of sudden, dramatic changes in equilibrium behavior, together with the difficulty of reversing these changes through increased enforcement, should concern policymakers.

Our analysis suggests a number of avenues for future research. First, although tax compliance in the United States (and in some other countries) is arguably in a mostly honest state, our models predict that, at some point, a slight reduction in enforcement efforts may lead to a tax evasion epidemic. However, our work does not identify the level of enforcement that would trigger a sudden, dramatic shift to noncompliance. In the agent-based model, the shift occurred at different audit rates across different initial conditions, so the distribution of individual characteristics in a society may affect the point at which this transition occurs. Further investigation can determine what factors (if any) one might use to predict the transition point.

Continuing with our epidemic analogy, one also might evaluate the extent to which the taxpaying population can be "inoculated" against evasion (in particular, what psychological or sociological mechanisms might be brought to bear and how effective such mechanisms might be) and the aggregate consequences of inoculating a proportion of the population (e.g., how much of the population must be protected to eliminate the possibility of an evasion epidemic). While in a different context, King's (2002) work examining the extent to which group membership reduces self-serving biases in auditors serves as an analogy. From an analytical perspective, the herd-immunity literature might provide insights (e.g., see Edelstein-Keshet 1988).

Third, although we find similar results based on two different modeling approaches, an experimental investigation could lend even more credence to our findings. Borrowing from the experimental literature in social psychology and accounting that creates social groups and norms in the laboratory (e.g., Smith et al. 1998; Schmader and Major 1999; King 2002), one might establish an experimental environment like the one our model describes to test its predictions.

Researchers might use the class of models we employ to address other questions in accounting where the behavior of others influences one's own behavior. In auditing, researchers often characterize reputation as arising from one's personal experience with an auditor (Mayhew et al. 2001). Using our approach, one could expand the notion of reputation to explicitly recognize that one's notion of an auditor's reputation probably arises both from one's own experience and from the experiences and beliefs of others. We speculate that recognition of the influence of others could lead to some unexpected insights, such as robustness of reputation in the face of frequent audit failures. Similarly, norms established in the profession, by competitors' behaviors and by sanctions, likely affect behaviors that auditors view as acceptable from an independence perspective. Explicit recognition of these influences could lead to new insights regarding auditor behavior.

Researchers might also develop influence models to better understand financial reporting behavior (e.g., earnings management). According to Healy and Wahlen (1999), future research on earnings management should identify limiting factors. One might model the behavior of other firms and reporting norms as limiting factors. For example, if competitors engage in earnings management, then the firm may feel pressure to follow suit. Reporting norms may play a countervailing role. Our approach could also yield new insights into financial analyst behavior by explicitly accounting for the influence that analysts exert on one another.

In a broader sense, we believe that the methods we employ have the potential to contribute to accounting research by overcoming the tendency to draw from a single discipline, such as psychology, economics, or finance. Both of our modeling approaches are amenable to integration of ideas from across the social sciences. For example, we draw the first principles for our models from social psychology and sociology. In our analysis, we derive equilibria (viz., economics) and evaluate dynamics using techniques drawn from biology. Other work using dynamic systems analysis and agent-based models in the social sciences has integrated research from sociology, political science, and even anthropology to gain new understanding about societies, markets, and individual behavior.

APPENDIX

Local Stability Analysis

To evaluate the stability of the equilibria in our system, we consider populations that are slightly out of equilibrium. Assume that for any $x \in [0, 1]$, $g(x)$ has a convergent Taylor series near x and let $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ be an equilibrium population. Then, we can express $g(x_1)$ as:

$$g(x_1) = g(\hat{x}_1) + g'(\hat{x}_1)(x_1 - \hat{x}_1) + O[(x_1 - \hat{x}_1)^2] \tag{14}$$

for x_1 sufficiently near $\hat{x}_1$. We can then derive rates of change for the deviations from any equilibrium as follows:

$$\begin{aligned}(x_1 - \hat{x}_1)' &= x_1' \\ &= [g(x_1) - rx_1]x_3 \\ &= (g(\hat{x}_1) - r\hat{x}_1 + [g'(\hat{x}_1) - r](x_1 - \hat{x}_1) + O[(x_1 - \hat{x}_1)^2])(x_3 - (\hat{x}_3 - \hat{x}_3)) \\ &= [g'(\hat{x}_1) - r]\hat{x}_3(x_1 - \hat{x}_1) + [g(\hat{x}_1) - r\hat{x}_1](x_3 - \hat{x}_3) \\ &\quad + [g'(\hat{x}_1) - r](x_1 - \hat{x}_1)(x_3 - \hat{x}_3) + O[(x_1 - \hat{x}_1)^2],\end{aligned}$$

where $O[(x_1 - \hat{x}_1)^2]$ indicates that all remaining terms in the expression are second and higher powers of $(x_1 - \hat{x}_1)$. The last step is justified because $[g(\hat{x}_1) - r\hat{x}_1]\hat{x}_3 = 0$ if $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ is an equilibrium point. A similar analysis for $x_3 - \hat{x}_3$ yields:

$$\begin{aligned}(x_3 - \hat{x}_3)' &= x_3' \\ &= \alpha - \alpha x_1 - \alpha x_3 - g(x_1)x_3 \\ &= \alpha - \alpha(x_1 - \hat{x}_1 + \hat{x}_1) - \alpha(x_3 - \hat{x}_3 + \hat{x}_3) - [g(\hat{x}_1) + g'(\hat{x}_1)(x_1 - \hat{x}_1) \\ &\quad + O[(x_1 - \hat{x}_1)^2]](x_3 - \hat{x}_3 + \hat{x}_3) \\ &= \alpha - \alpha\hat{x}_1 - \alpha\hat{x}_3 - g(\hat{x}_1)\hat{x}_3 - \alpha(x_1 - \hat{x}_1) - [g(\hat{x}_1) + \alpha](x_3 - \hat{x}_3) \\ &\quad - g'(\hat{x}_1)(x_1 - \hat{x}_1)(x_3 - \hat{x}_3) + O[(x_1 - \hat{x}_1)^2] \\ &= -\alpha(x_1 - \hat{x}_1) - [g(\hat{x}_1) + \alpha](x_3 - \hat{x}_3) - g'(\hat{x}_1)(x_1 - \hat{x}_1)x_3 \\ &\quad + O[(x_1 - \hat{x}_1)^2].\end{aligned}$$

Again, the last step is justified because $\alpha - \alpha\hat{x}_1 - \alpha\hat{x}_3 - g(\hat{x}_1)\hat{x}_3 = 0$ at equilibrium. We combine the equations for $(x_1 - \hat{x}_1)'$ and $(x_3 - \hat{x}_3)'$ and write them in matrix form, leading to the following linearized system of equations near the equilibrium (we drop the higher-order term):

$$\begin{aligned}\begin{pmatrix} x_1 - \hat{x}_1 \\ x_3 - \hat{x}_3 \end{pmatrix}' &= \begin{pmatrix} [g'(\hat{x}_1) - r]\hat{x}_3 & g(\hat{x}_1) - r\hat{x}_1 \\ -\alpha & -[\alpha + g(\hat{x}_1)] \end{pmatrix} \begin{pmatrix} x_1 - \hat{x}_1 \\ x_3 - \hat{x}_3 \end{pmatrix} \\ &\quad + \begin{pmatrix} g'(\hat{x}_1) - r \\ -g'(\hat{x}_1) \end{pmatrix} (x_1 - \hat{x}_1)(x_3 - \hat{x}_3) + \begin{pmatrix} 0 \\ -g'(\hat{x}_1) \end{pmatrix} \hat{x}_3. \end{aligned} \tag{15}$$

Now we examine the stability of the linearized system near the honest equilibrium. In particular, near a type (i) or honest equilibrium, the linearized system in Equation (15) takes the form:

$$\begin{pmatrix} x_1 - 1 \\ x_3 \end{pmatrix}' = \begin{pmatrix} 0 & g(1) - r \\ -\alpha & -[\alpha + g(1)] \end{pmatrix} \begin{pmatrix} x_1 - 1 \\ x_3 \end{pmatrix} + \begin{pmatrix} g'(1) - r \\ -g'(1) \end{pmatrix} (x_1 - 1)x_3.$$

The eigenvalues of the 2 × 2 matrix for the linear part of the system characterize the stability of the equilibrium at (1, 0, 0). Specifically, by the local stability theorem, (1, 0, 0) is stable if and only if both eigenvalues have negative real parts. It is unstable if either real part is positive. Hence, to prove Proposition 3, we must first find a λ to satisfy:

$$\det \begin{vmatrix} -\lambda & g(1) - r \\ -\alpha & -\lambda - [\alpha + g(1)] \end{vmatrix} = 0.$$

This gives the equation:

$$\begin{aligned} 0 &= \lambda(\lambda + [\alpha + g(1)]) + \alpha[g(1) - r] \\ &= \lambda^2 + [\alpha + g(1)]\lambda + \alpha[g(1) - r]. \end{aligned}$$

Solving the resulting quadratic equation for λ, we find:

$$\lambda = -\left(\frac{\alpha + g(1)}{2}\right)\left(1 \pm \sqrt{1 - \frac{4\alpha(g(1) - r)}{(\alpha + g(1))^2}}\right). \tag{16}$$

The roots are complex if and only if:

$$1 - \frac{4\alpha(g(1) - r)}{(\alpha + g(1))^2} < 0. \tag{17}$$

We rewrite this condition as:

$$[\alpha - g(1)]^2 < -4\alpha r. \tag{18}$$

Because α and r are positive, Equation (18) is never satisfied. The eigenvalues are real. The eigenvalue corresponding to $1 + \sqrt{\ldots}$ is negative. We determine stability using the sign of the other root. That is, (1, 0, 0) is unstable if and only if:

$$\begin{aligned} 1 - \sqrt{1 - \frac{4\alpha(g(1) - r)}{(\alpha + g(1))^2}} &< 0 \\ 1 &< \sqrt{1 - \frac{4\alpha(g(1) - r)}{(\alpha + g(1))^2}} \\ 0 &< -\frac{4\alpha(g(1) - r)}{(\alpha + g(1))^2} \\ 0 &< r - g(1) \\ g(1) &< r. \end{aligned}$$

Conversely, (1, 0, 0) is stable if and only if $g(1) > r$.

Extending our analysis to the linearized system near a type (ii) or mixed equilibrium, we obtain:

$$\begin{pmatrix} x_1 - \hat{x}_1 \\ x_3 - \hat{x}_3 \end{pmatrix}' = \begin{pmatrix} [g'(\hat{x}_1) - r]\hat{x}_3 & 0 \\ -\alpha & -[\alpha + g(\hat{x}_1)] \end{pmatrix} \begin{pmatrix} x_1 - \hat{x}_1 \\ x_3 - \hat{x}_3 \end{pmatrix} + \begin{pmatrix} g'(\hat{x}_1) - r \\ -g(\hat{x}_1) \end{pmatrix} (x_1 - \hat{x}_1)(x_3 - \hat{x}_3) + \begin{pmatrix} 0 \\ -g'(\hat{x}_1) \end{pmatrix} \hat{x}_3. \tag{19}$$

The eigenvalues of the linear part are:

$$\lambda_1 = -[\alpha + g(\hat{x}_1)] \quad \text{and} \quad \lambda_2 = [g'(\hat{x}_1) - r]\hat{x}_3.$$

Because λ_1 is necessarily negative, we determine stability of the equilibrium by the sign of $g'(\hat{x}_1) - r$. This leads to Proposition 4.

REFERENCES

Albin, P., and D. K. Foley. 1992. Decentralized, dispersed exchange without an auctioneer: A simulation study. *Journal of Economic Behavior and Organization* 18 (1): 27–51.

Alm, J. 1991. A perspective on the experimental analysis of taxpayer reporting. *The Accounting Review* 66 (3): 577–593.

———, I. Sanchez, and A. de Juan. 1995. Economic and noneconomic factors in tax compliance. *Kyklos* 48 (1): 3–18.

Arnol'd, V. 1991. *Ordinary Differential Equations*. 3rd edition. New York, NY: Springer-Verlag.

Arthur, W. B., J. H. Holland, B. LeBaron, R. Palmer, and P. Tayler. 1997. Asset pricing under endogenous expectations in an artificial stock market. In *The Economy as an Evolving Complex System II*, edited by W. B. Arthur, S. N. Durlauf, and D. A. Lane, Vol. 27 of Santa Fe Institute Studies in the Sciences of Complexity, 1–44. Reading, MA: Addison-Wesley.

Axelrod, R. 1997. *The Complexity of Cooperation*. Princeton, NJ: Princeton University Press.

Axtell, R. 2000. Why agents? On the varied motivations for agent computing in the social sciences. In *Proceedings of the Workshop on Agent Simulation: Applications, Models, and Tools*, edited by C. M. Macal, and D. Sallach, 3–24. Chicago, IL: University of Chicago, Argonne National Laboratory.

Bloomfield, R. 1995. Strategic dependence and inherent risk assessments. *The Accounting Review* 70 (1): 71–90.

———. 1997. Strategic dependence and the assessment of fraud risk: A laboratory study. *The Accounting Review* 72 (4): 517–538.

Carley, K. 1991. A theory of group stability. *American Sociological Review* 56 (3): 331–354.

Cooter, R. 1998. Expressive law and economics. *Journal of Legal Studies* 27 (2): 585–607.

Crist, T., and J. Haefner. 1994. Spatial model of movement and foraging in harvester ants (*Pogonomyrmex*) II: The roles of environment and seed dispersion. *Journal of Theoretical Biology* 166 (3): 315–323.

Danielson, P., ed. 1996. *Modelling Rationality, Morality, and Evolution*. New York, NY: Oxford University Press.

de Juan, A., M. Lasheras, and R. Mayo. 1994. Voluntary tax compliance behavior of Spanish income taxpayers. *Public Finance* 49 (Supplement): 90–105.

Edelstein-Keshet, L. 1988. *Mathematical Models in Biology*. Berkhäuser Mathematical Series. New York, NY: Random House.

Epstein, J., and R. Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Cambridge, MA: MIT Press.

———. 1997. *Nonlinear Dynamics, Mathematical Biology, and Social Science*. Vol. 4 of Santa Fe Institute Lecture Notes. Reading, MA: Addison-Wesley.

Erard, B., and J. Feinstein. 1994. Honesty and evasion in the tax compliance game. *The Rand Journal of Economics* 25 (1): 1–19.

Gaylord, R., and J. Davis. 1999. Modeling nonspatial socioeconomic interactions: An agent-based approach. *Mathematica in Education and Research* 8 (2): 30–33.

Geeroms, H., and H. Wilmots. 1985. An empirical model of tax evasion and tax avoidance. *Public Finance* 40 (2): 190–209.

Glance, N. S., and B. A. Huberman. 1993. The outbreak of cooperation. *Journal of Mathematical Sociology* 17 (4): 281–302.

Grasmick, H., and D. Green. 1980. Legal punishment, social disapproval, and internalization as inhibitors of illegal behavior. *Journal of Criminal Law and Criminology* 71 (3): 325–335.

———, and W. Scott. 1982. Tax evasion and mechanisms of social control: A comparison with grand and petty theft. *Journal of Economic Psychology* 2 (3): 213–230.

Haefner, J., and T. Crist. 1994. Spatial model of movement and foraging in harvester ants (Pogonomyrmex) I: The roles of memory and communication. *Journal of Theoretical Biology* 166 (3): 299–313.

Hamblin, R., R. Jacobsen, and J. Miller. 1973. *A Mathematical Theory of Social Change*. New York, NY: Wiley.

Healy, P. M., and J. M. Wahlen. 1999. A review of the earnings management literature and its implications for standards setting. *Accounting Horizons* 13 (4): 365–383.

Hite, P. 1988. The effect of peer reporting behavior on taxpayer compliance. *The Journal of the American Taxation Association* 10 (2): 47–64.

Kahneman, D., and A. Tversky. 1973. On the psychology of prediction. *Psychological Bulletin* 80 (4): 237–251.

———, and ———. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2): 263–291.

Kaplan, S., and P. Reckers. 1985. A study of tax evasion judgments. *National Tax Journal* 38 (1): 97–102.

Kermack, W., and A. McKendrick. 1927. Contributions to the mathematical theory of epidemics. *Proceedings of the Royal Statistical Society* (Series A) 115: 700–721.

King, R. 2002. An experimental investigation of self-serving biases in an auditing trust game: The effect of group affiliation. *The Accounting Review* 77 (2): 265–284.

Kirman, A. 1992. Whom or what does the representative agent represent? *Journal of Economic Perspectives* 6 (2): 117–136.

Kohler, T. A., J. Kresl, C. Van West, E. Carr, and R. Wilshusen. 1999. *Dynamics in Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes*. Santa Fe, NM: Santa Fe Institute.

Kollman, K., J. Miller, and S. Page. 1997. Political institutions and sorting in a Tiebout model. *American Economic Review* 87 (5): 977–992.

Lerner, M. 1998. The two forms of belief in a just world. In *Responses to Victimizations and Belief in a Just World*, edited by L. Montada, and M. Lerner, 247–269. New York, NY: Plenum.

Lewis, A. 1979. An empirical assessment of tax mentality. *Public Finance* 43 (2): 245–257.

Lindbeck, A., S. Nyberg, and J. Weibull. 1999. Social norms and economic incentives in the welfare state. *Quarterly Journal of Economics* 114 (1): 1–35.

Marengo, L., and H. Tordjman. 1996. Speculation, heterogeneity, and learning: A simulation model of exchange rate dynamics. *Kyklos* 49 (3): 407–438.

Marimon, R., E. McGrattan, and T. J. Sargent. 1990. Money as a medium of exchange in an economy with artificially intelligent agents. *Journal of Economic Dynamics and Control* 14 (2): 329–373.

Mayhew, B., J. Schatzberg, and G. Sevcik. 2001. The effect of accounting uncertainty and auditor reputation on auditor objectivity. *Auditing: A Journal of Practice & Theory* 20 (2): 49–70.

McAdams, R. 1997. The origin, development, and regulation of norms. *Michigan Law Review* 96 (2): 338–433.

Murray, J. 1989. *Mathematical Biology*. New York, NY: Springer-Verlag.

Nisbett, R. E., E. Borgida, R. Crandall, and H. Reed. 1976. Popular induction: Information is not always informative. In *Cognition and Social Behavior*, edited by J. S. Carroll, and J. W. Payne, Vol. 2, 227–236. Hillsdale, NJ: Erlbaum.

Page, S., and L. Hong. 1996. Problem solving by teams of heterogeneous agents. Unpublished manuscript, Pasadena, CA, California Institute of Technology, Division of Humanities and Social Sciences.

Plous, S. 1993. *The Psychology of Judgment and Decision Making*. New York, NY: McGraw-Hill.

Schelling, T. 1969. Models of segregation. *American Economic Review* 59 (2): 488–493.

———. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1 (2): 143–186.

———. 1978. *Micromotives and Macrobehavior*. New York, NY: Norton.

Schepanski, A., and D. Kelsey. 1990. Testing for framing effects in taxpayer compliance decisions. *The Journal of the American Taxation Association* 12 (1): 60–77.

Schmader, T., and B. Major. 1999. The impact of ingroup versus outgroup performance on personal values. *Journal of Experimental Social Psychology* 35 (1): 47–67.

Scott, W., and H. Grasmick. 1981. Deterrence and income tax cheating. *Journal of Applied Behavioral Science* 17 (3): 395–408.

Smith, H. J., T. R. Tyler, Y. J. Huo, D. J. Ortiz, and E. A. Lind. 1998. The self-relevant implications of the group-value model: Group membership, self-worth, and treatment quality. *Journal of Experimental Social Psychology* 34 (5): 470–493.

Spicer, M., and S. Lundstedt. 1976. Understanding tax evasion. *Public Finance* 31 (2): 295–305.

———, and L. Becker. 1980. Fiscal inequity and tax evasion: An experimental approach. *National Tax Journal* 33 (2): 171–175.

Tversky, A., and D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5 (2): 207–232.

Tyler, T. R. 1990. *Why People Obey the Law*. New Haven, CT: Yale University Press.

Vogel, J. 1974. Taxation and public opinion in Sweden: An interpretation of recent survey data. *National Tax Journal* 27 (4): 499–513.

Waltman, P. 1974. *Deterministic Threshold Models in the Theory of Epidemics*. Lecture Notes in Biomathematics, Vol. 1. New York, NY: Springer-Verlag.

Weibull, J. W. 1997. *Evolutionary Game Theory*. Cambridge, MA: MIT Press.

Weisbuch, G., A. Kirman, and D. Herreiner. 1998. Market organisation. In *Advances in Self-Organisation and Evolutionary Economics*, edited by J. Lesouren, and A. Orlean, 160–192. London, U.K.: Economica.

Weisman, J. 2001. IRS audits at record low: Officials cite decade-long staff decrease, old computers, law defending taxpayers. *USA Today* (February 16): A1.

Westat, Inc. 1980. *Individual Income Tax Compliance Factors Study: Qualitative Research Results*. Rockville, MD: Westat.

White, R., P. Harrison, and A. Harrell. 1993. The impact of income tax withholding on taxpayer compliance: Further empirical evidence. *The Journal of the American Taxation Association* 15 (2): 63–78.

Witte, A., and D. Woodbury. 1985. The effects of tax laws and tax administration on tax compliance: The case of the U.S. individual income tax. *National Tax Journal* 38 (1): 1–13.

Yankelovich, Skelly, and White, Inc. 1984. *Taxpayer Attitudes Survey: Final Report*. New York, NY: Internal Revenue Service, Public Affairs Division.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 71–93

Capital Budgeting and Managerial Compensation: Incentive and Retention Effects

Sunil Dutta
University of California, Berkeley

ABSTRACT: This paper considers an agency model in which a firm's manager receives private information about an investment project. The manager has unique skills that are essential for implementing the project, and he can pursue the project inside the firm or as an outside venture on his own. The firm's owner thus faces a potential managerial retention problem, where the severity of the retention problem depends on the project's outside viability. My analysis shows that if the managerial retention problem is not too severe, the owner can delegate the investment decision to the manager and use a residual-income-based bonus contract to give the manager incentives to work hard and make appropriate investment decisions. If the retention problem is severe, however, then the owner must use an option-based compensation contract to retain the manager and provide him with appropriate incentives. I also establish that as the managerial retention problem becomes more severe, the owner reduces the rate of return, or hurdle rate, required to approve the investment project. These results predict that new-economy firms, in which managerial expertise is critical and yet mobile, are more likely to (1) include stock options in their managers' compensation contracts, and (2) apply lower hurdle rates for approving capital investments.

Keywords: *capital investments; stock options; managerial retention; managerial incentives.*

I. INTRODUCTION

Corporate managers often identify potentially profitable investment projects and develop the expertise or human capital necessary to successfully execute these projects. The speed of technological change in recent years has increased the importance of specialized expertise or human capital, which is now much more mobile than in

---

I thank Raffi Indjejikian (associate editor) and two anonymous reviewers for helpful comments and suggestions. I also acknowledge the useful comments of Tim Baldenius, Ron Dye, Frank Gigler, Stefan Reichelstein, Brett Trueman, and the seminar participants at Northwestern University.

Submitted January 2001
Accepted August 2002

the past.[1] As a result, managers can more easily walk away with the relevant information and expertise, and pursue the projects on their own or with other employers. Although firms attempt to restrict their employees from exploiting information and knowledge acquired during their employment, anecdotal and survey evidence suggest that such restrictions are ineffective. Bhide (2000) reports that founders of more than 70 percent of the Inc. 500 firms admitted that they developed and marketed ideas identified in their previous employment. Protection of new investment opportunities is likely to become increasingly challenging for new-economy firms that rely heavily on managerial human capital. Indeed, Zingales (2000) argues that retaining investment opportunities and managerial expertise is one of the most important corporate governance issues for these firms.[2]

Consequently, in structuring executive compensation plans, firms not only must consider how to motivate their managers to work hard and make appropriate investment decisions, but also must ensure that they retain valuable managerial human capital. Although a large and growing literature examines how firms can design compensation plans to give their managers desirable effort and investment incentives, this research gives scant attention to the steps firms must take to retain their managers. This paper provides a formal analysis of the managerial retention problem in a principal-agent setting.

In this paper, I model a manager who receives private information about a potential investment opportunity that he could undertake with his firm. The manager also exerts an unobservable level of effort to increase the firm's cash flows. In addition, the manager has unique expertise that is essential to successfully complete the project. The manager can pursue the project inside the firm or as an outside venture on his own. My analysis characterizes the firm's investment decision rule and its choice of a managerial compensation contract in two different organizational settings: (1) a *centralized* capital-budgeting setting in which the firm's owner approves the investment project herself; and (2) a *delegated* investment setting in which the owner delegates the investment decision to the manager.

In the centralized capital-budgeting setting, the manager reports his private information about the investment project to the owner, who then decides whether to approve the project. In my model, the manager has an incentive to understate the project's return because it allows him to exert minimal effort and then ascribe the resulting poor performance to low investment returns. To counteract this incentive for misreporting, the owner must pay the manager a compensation premium, which I call the *incentive premium.*

Since the manager can benefit from misrepresenting the project's profitability only if the owner authorizes the investment, the owner pays the incentive premium to the manager only when she approves the project. The owner therefore declines marginally profitable projects, which increases her net profits by reducing the compensation she must pay the manager. Consistent with Antle and Eppen (1985) and others, I thus find that the owner optimally rations capital—that is, she requires a rate of return (or *hurdle rate*) that exceeds her cost of capital. The owner chooses the hurdle rate so that the marginal "benefit" of underinvestment (i.e., the reduced compensation premium she pays the manager) is equal to the marginal opportunity cost (i.e., the positive net present value forgone by declining the project).

To ensure that the manager finds it in his self-interest to stay with the firm and execute the project inside the firm, the owner also may have to pay him an additional compensation premium, which I call the *retention premium.* Although the owner can reduce the incentive portion of the manager's compensation premium by rejecting marginally profitable projects, her project acceptance rule does not affect the compensation premium necessary to retain the manager. Consequently, when the project's outside viability increases, and hence the managerial retention problem becomes more severe, the owner's marginal benefit of underinvestment declines. As a result, the owner finds it optimal to introduce *less* underinvestment—that is, she requires a lower hurdle rate for project approval. This result provides a potentially testable prediction regarding cross-sectional variation in firms' choices of hurdle rates for approval of capital investment projects. In particular, it suggests that firms will set lower hurdle rates for project approval when: (1) their investment projects require specific managerial expertise, or human capital; and (2) managers can execute these projects in outside ventures.

---

[1] For an excellent review of the changing nature of business enterprises, see Rajan and Zingales (2000).

[2] Ittner et al. (2001) also argue that managerial retention becomes a more important issue when managers can walk away with information and technology.

In the delegated investment setting that I consider, the owner assigns the investment decision rights to the manager, who does *not* communicate his information to the owner. The delegated investment scenario is important because in practice corporate managers frequently make investment decisions for their firms or divisions. In order to align their managers' incentives with shareholders' goals, decentralized firms typically evaluate their managers' performance on measures such as share price, return on assets, income, and residual income. In the delegated investment setting, the manager need not communicate all of the project details to the owner. However, because the owner no longer dictates the firm's investment choice, the owner must give the manager appropriate incentives to make desirable investments.

We know from the revelation principle that if communication is costless, a decentralized organization's performance can never dominate a centralized organization's performance. The important question in the delegated setting is whether the owner can design an incentive plan (i.e., a performance measure and compensation rule) for the manager so that the owner attains the same expected payoff under delegation as she would earn under centralized capital budgeting. I refer to such a managerial compensation contract as *optimal*, in that it gives optimal effort and investment incentives.[3]

My analysis establishes that if the managerial retention problem is not too severe, the owner can use a bonus contract based on residual income to give optimal effort and investment incentives. The capital charge rate used in the residual income measure, however, must equal the owner's optimal hurdle rate, not her actual cost of capital.[4] In contrast, I find that when the managerial retention problem is severe, a bonus contract based on residual income can no longer provide the manager with optimal incentives. Instead, the owner must include an option-like component in the manager's compensation contract—that is, a bonus paid only if the manager's performance exceeds some minimum threshold. My paper thus identifies conditions under which managerial compensation packages will include option-like components. In particular, my results predict that new-economy, knowledge-intensive firms (in which managerial expertise is critical as well as mobile) are more likely to include option-like components in their managers' compensation contracts.

This prediction is consistent with empirical and anecdotal evidence that stock options play a more prominent role in the compensation of executives of new-economy, knowledge-intensive firms than in the compensation of executives of traditional firms (see, for instance,

---

[3] If the owner can identify such a compensation contract, and if it is costly to communicate the project information, she will *strictly* prefer delegation to centralization.

[4] This is consistent with Christensen et al. (2002) and Dutta and Reichelstein (2002), who also show that the capital charge rate in agency settings generally differs from the firm's cost of capital.

Anderson et al. 2000; Core and Guay 2001; Ittner et al. 2001).[5] Furthermore, in my model, the owner uses the option component not so much to provide effort and investment incentives as to retain valuable managerial human capital and profitable investment projects inside the firm. This result is consistent with Ittner et al. (2001), who find that option grants help firms attract and retain employees.[6]

In addition to providing a theoretical rationale for the increasing importance of option-like management compensation contracts, my paper contributes to two other streams of the accounting literature. The first stream analyzes the effect of asymmetric information on firms' capital investment decisions. Antle and Eppen (1985) show that an owner will ration capital in order to reduce a privately informed manager's compensation premium. My paper extends their analysis by examining how the managerial retention problem affects the owner's project approval rule. My results show that as the managerial retention problem becomes more severe, the owner requires a lower hurdle rate for project approval. A more recent stream of the literature concludes that firms can use residual-income-based compensation plans to provide their managers with desirable investment incentives (Rogerson 1997; Reichelstein 1997). My analysis extends this literature by establishing that new-economy, knowledge-intensive firms may have to supplement residual-income-based managerial compensation contracts with option-like components.

The remainder of the paper proceeds as follows. Section II describes the economic setting. Section III characterizes the owner's optimal project approval rule under centralized capital budgeting. Section IV examines the delegated investment setting and characterizes optimal managerial compensation contracts. Section V concludes the paper.

II. THE MODEL

I model a one-period principal-agent relationship between a firm's risk-neutral owner (principal) and its risk-neutral manager (agent). The period begins at date 0 and ends at date 1. At the beginning of the period, the manager receives information about a potentially profitable investment project. Undertaking this project requires a cash outlay of k dollars at date 0, and generates cash inflow of $(1 + r) \cdot k$ dollars at date 1. Let c denote the firm's cost of capital. Then the project's date 0 net present value is $k \cdot (r - c) \cdot (1 + c)^{-1}$, where $(1 + c)^{-1}$ is the discount factor corresponding to the firm's cost of capital c. While both the owner and the manager know the investment cost k, only the manager observes the project's rate of return r. The owner's prior beliefs regarding the project's rate of return r are represented by a distribution F(r) with positive density f(r) on the interval $[\underline{r}, \bar{r}]$.

The end-of-period cash inflow x is:

$$x = y + (1 + r) \cdot k \cdot I + a, \tag{1}$$

where y denotes the commonly known cash returns from the firm's existing assets. Since y is a known constant, I normalize it to zero without loss of generality. The indicator variable $I \in \{0,1\}$ denotes the publicly observed investment choice, which equals 1 when the firm undertakes the project, and 0 otherwise. Finally, a denotes the manager's unobservable effort choice, which increases the firm's cash flows. Note that the managerial effort improves the firm's cash flows regardless of the investment choice I. The interpretation of this assumption is that managerial effort enhances the cash returns from the firm's existing assets, and the productivity of the managerial effort does not vary with the investment choice.

---

[5] Ely (1991) also reports empirical evidence that managerial incentive plans differ systematically across different industries.

[6] In my one-period model, I assume the manager has precontract private information. Therefore, one might argue that the paper addresses a managerial hiring or attraction problem, rather than a managerial retention problem. As explained in Section II, however, my model permits both of these interpretations.

The owner observes the aggregate cash flow in Expression (1), but not its individual components related to the project and to the manager's action. The owner's inability to disaggregate the components of the cash flows, combined with the manager's private information regarding the project's rate of return r, creates an agency problem that prevents the owner from achieving the first-best outcome. Specifically, in my model, the manager has a natural incentive to shirk and then ascribe the resulting poor performance to low investment returns.

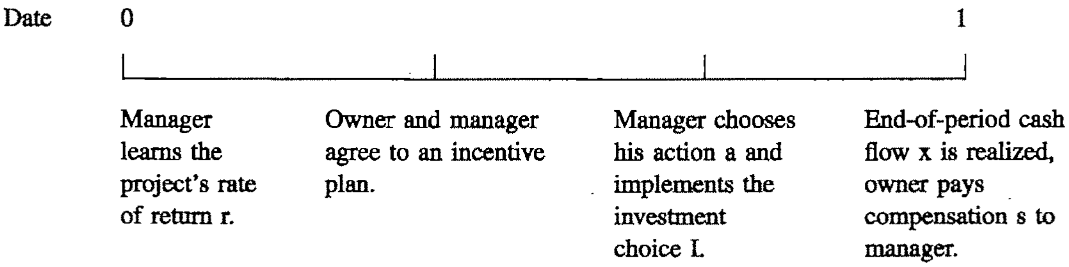
The key feature I seek to model is the notion that the manager can walk away with the relevant information and expertise about the investment project, and hence the firm faces a potential managerial retention problem. To model a managerial retention threat, I assume that the manager (1) has unique skills that are necessary to complete the project, and (2) can undertake the investment project on his own in an outside venture. If he undertakes the project outside the firm, then it yields a rate of return of $\delta \cdot r$, where $\delta \in [0,1)$ is a commonly known parameter.

I assume that the project is more profitable inside than outside the firm (i.e., $\delta < 1$) because there are synergies between the firm's existing assets and the new investment opportunity. As δ decreases, the project's rate of return in the outside venture declines. When $\delta = 0$, the project is entirely firm-specific because it yields no payoff outside the firm. On the other hand, as δ approaches 1, the project's profitability outside the firm approaches its profitability inside the firm (i.e., the new project is entirely independent of the firm's existing assets).

The manager chooses effort a from the interval $[0, \bar{a}]$ at a personal cost of $v \cdot a$, with $v \in (0,1)$.[7] The risk-neutral manager's utility (net payoff) is then given by:

$$u = s - v \cdot a$$

where s denotes his compensation. The time line below depicts the sequence of events in the model.


Date
0
1
Manager learns the project's rate of return r.
Owner and manager agree to an incentive plan.
Manager chooses his action a and implements the investment choice I.
End-of-period cash flow x is realized, owner pays compensation s to manager.

The manager has precontract private information about the project's rate of return r. As a consequence, the manager will accept an incentive plan only if it provides him with at least his reservation utility for each value of r. One interpretation of the precontract private information setting is that the incumbent manager's contract from the prior period (not modeled here) has expired. The owner now seeks to design a new compensation plan to *retain* the manager and to give him appropriate incentives to exert effort.

---

[7] As explained in footnote 19, my results continue to hold even when the manager's cost of effort is nonlinear.

An alternative interpretation is that the firm wishes to hire a new entrepreneur/manager who has relevant information and expertise about a potentially profitable investment opportunity. The entrepreneur can pursue the investment opportunity on his own or join the firm and implement it inside the firm. Under this interpretation, the owner designs a compensation package to *attract* the entrepreneur/manager and give him incentives to exert effort.[8] Given these alternative but analytically equivalent interpretations of the model, I use the terms *retention* and *attraction* interchangeably in the rest of the paper.

III. CENTRALIZED CAPITAL BUDGETING

In centralized firms, divisional managers (who often have superior information about the investment opportunities in their divisions) submit proposals for new investments to their corporate headquarters, which then determines whether to approve the proposals. I refer to this form of investment decision making as centralized capital budgeting. This section characterizes the owner's optimal investment choice in a centralized capital budgeting setting in which the owner makes the investment decision based on the manager's report about the project's profitability.

The manager submits a report $\hat{r}$ about the project's rate of return to the owner after the owner has committed to an incentive plan for the manager. Given the revelation principle, I consider only incentive plans that induce the manager to reveal his information truthfully. A managerial incentive plan in the centralized capital budgeting setting specifies the manager's compensation $s(\hat{r})$, the required performance $x(\hat{r})$ (measured in terms of the end-of-period cash flow), and the investment decision rule $I(\hat{r})$, all as functions of the manager's report $\hat{r}$.

When the project's true rate of return is r but the manager reports $\hat{r}$, let $a(\hat{r},r)$ denote the minimum amount of effort that the manager must expend to generate cash flow consistent with his report $\hat{r}$. That is:

$$a(\hat{r},r) \equiv x(\hat{r}) - (1 + r) \cdot k \cdot I(\hat{r}). \tag{2}$$

Contingent on the reported return of $\hat{r}$ and the true return of r, the manager's net payoff becomes:

$$u(\hat{r},r) \equiv s(\hat{r}) - v \cdot a(\hat{r},r). \tag{3}$$

Let u(r) denote the manager's net payoff when he reports his private information truthfully; that is:

$$u(r) \equiv u(r,r). \tag{4}$$

The risk-neutral owner chooses managerial incentive plan $\{s(r), x(r), a(r)\}$ to maximize the expected value of her net profit. The following program describes the owner's optimization problem:

---

[8] Alternatively, one can assume that the manager receives postcontract private information, but he can quit after he learns r. Since the owner must again ensure the manager's participation for each value of r, the postcontract private information setting is equivalent to the precontract private information scenario considered in the paper.

$$\max_{I(r),s(r),x(r)} \int_{\underline{r}}^{\bar{r}} [x(r) - s(r) - (1 + c) \cdot k \cdot I(r)]\, f(r)\, dr$$

subject to:

$$u(r) \geq 0 \qquad \text{for all } r \qquad \text{(PC)}$$

$$u(r) \geq k \cdot (\delta \cdot r - c) \qquad \text{for all } r \qquad \text{(RC)}$$

$$u(r) \geq u(\hat{r}, r) \qquad \text{for all } \hat{r} \text{ and } r. \qquad \text{(IC)}$$

The objective function in the above program reflects the expected value of the owner's net payoff at date 1. The participation constraint (PC) ensures that the manager earns at least his reservation utility, which I have normalized to zero without loss of generality. The retention constraint (RC) guarantees that the manager will stay with the firm rather than quit and pursue the project outside the firm.[9] The incentive compatibility constraint (IC) ensures that the manager will find it in his self-interest to report his private information truthfully.

Optimal Investment in the Absence of a Managerial Retention Problem

Before solving the owner's optimization problem in the above optimization program, I consider a benchmark problem in which the project has no value outside the firm (i.e., $\delta = 0$), and therefore the owner does not face any managerial retention problem. Because the retention constraint no longer applies, the optimization problem reduces to a standard adverse-selection problem. To examine the optimal solution to the benchmark problem, consider first a symmetric information setting in which the owner also observes the project's rate of return r. In this first-best setting, the owner approves the investment project whenever its net present value is positive and rejects it when the net present value is negative. That is, $I(r) = 1$ if $r \geq c$ and $I(r) = 0$ if $r < c$.

Since the manager's marginal cost of effort v is less than the marginal product of effort (normalized to 1), the owner will optimally seek to induce the maximum managerial effort, $\bar{a}$. A simple "forcing" compensation contract that requires the manager to deliver a cash flow of $\bar{a} + (1 + r) \cdot k \cdot I(r)$ for a fixed compensation of $v \cdot \bar{a}$ will yield the desired outcome. Therefore, if the owner and the manager have symmetric information about the project's profitability, then the manager will earn his reservation payoff for all values of r.

In my model, however, only the manager knows the project's rate of return r. As a consequence, the owner must design the managerial incentive plan so that it induces the manager to report the project's rate of return truthfully. Note that in the absence of a hidden action problem, the manager would have no incentive to misrepresent his private information, because he is indifferent toward all investment policies.[10] Because of the underlying hidden-effort problem, however, the manager has an incentive to understate the project's rate of return. By understating the project's profitability, the manager can deliver the required cash flows with less exertion of personally costly effort.[11]

---

[9] To see this, note that the gross payoff from implementing the project in the outside venture is $k \cdot (1 + \delta \cdot r)$, and the date 1 compounded value of the investment cost is $k \cdot (1 + c)$. Hence, the right-hand side of constraint (RC), $k \cdot (\delta \cdot r - c)$, denotes the project's net payoff outside the firm.

[10] In contrast, the "empire-building" literature assumes that managers derive private benefits from investments (Harris and Raviv 1996; Bernardo et al. 2001; Baldenius 2001; Lambert 2001).

[11] Antle and Eppen (1985) and Antle and Fellingham (1990) consider models in which the owner does not observe the amount of capital investment, and hence the manager can consume any funds in excess of the amount invested. In their models, therefore, nontrivial incentive problems exist even without any hidden action problem.

Because the manager's marginal cost of effort is constant, the owner will seek to induce either the minimum or maximum level of managerial effort. If the owner sought only the minimum level of managerial effort, then the manager would have no incentive to misreport his private information. To avoid this trivial case, I assume that the marginal return from managerial effort is sufficiently large (relative to the marginal cost of effort) that the owner prefers the maximum level of managerial effort.[12] As shown in the Appendix, incentive constraint (IC) then requires that:

$$u(r) = v \cdot k \cdot \int_{\underline{r}}^{r} I(z)\, dz \quad \text{for each } r. \tag{5}$$

The condition in Equation (5) implies that the manager's net payoff must increase at the rate of $v \cdot k$ in the project's profitability (in the range of states in which the owner approves the project—that is, when $I(r) = 1$). Combined with the manager's participation requirement that $u(r) \geq 0$, this condition implies that if the owner approves the project, the manager will earn more than his reservation utility (i.e., the manager will earn a compensation premium).

To understand why, suppose the owner pays the manager only for his cost of effort and pays no premium. Suppose further that the project's net present value is positive (i.e., $r > c$). If the manager reports his private information truthfully (i.e., $\hat{r} = r$), then the owner would require him to undertake the project and generate a cash flow of $\bar{a} + (1 + r) \cdot k$. To deliver the required cash flow, the manager would have to exert $\bar{a}$ units of effort. The manager, however, could do better by misreporting the project's rate of return. In particular, if the manager understates the project's rate of return by ε, then he could deliver the required cash flow (i.e., $\bar{a} + (1 + r - \varepsilon) \cdot k$) by contributing only $\bar{a} - k \cdot \varepsilon$ units of effort, and thus save $v \cdot k \cdot \varepsilon$ in his cost of effort.

To counteract the manager's incentives to understate the project's profitability, the owner must design the manager's incentive plan so that if the manager underreports the project's rate of return, his resulting savings from exerting less effort are exactly offset by lower compensation. Any incentive plan that satisfies Condition (5) has this property because the manager's compensation increases in his report at the rate of $v \cdot k$.

Using Expression (5) to solve for the compensation payment $s(r)$, the owner's optimization problem simplifies to the following unconstrained maximization problem:

$$\max_{I(r)} \int_{\underline{r}}^{\bar{r}} I(r) \cdot (r - c - v \cdot H(r))\, f(r)\, dr,$$

where $H(r) \equiv [1 - F(r)]/f(r)$ denotes the inverse hazard rate.[13] Pointwise optimization yields the following result:

Observation 1: When the project has no value outside the firm (i.e., when $\delta = 0$), the owner will approve the investment project if and only if its return r exceeds a hurdle rate r° that solves:

---

[12] As shown in the Appendix, a sufficient condition is that $\bar{a} \cdot (1 - v) \geq v \cdot k \cdot H(c)$, where $H(c) = [1 - F(c)] \cdot [f(c)]^{-1}$ denotes the inverse hazard rate.

[13] I make the standard assumption that the inverse hazard rate is decreasing. Many distributions have this property including the uniform, the normal, the exponential, and any other distribution with increasing density.

$$r^{o} = c + v \cdot \frac{1 - F(r^{o})}{f(r^{o})}. \tag{6}$$

Proof: All proofs are in the Appendix.

Consistent with Antle and Eppen (1985), Observation 1 establishes that the owner approves the investment project only if its return exceeds a hurdle rate r^{o}. Furthermore, the owner's choice of hurdle rate r^{o} exceeds her true cost of capital c—that is, the owner optimally chooses to underinvest. To understand the intuition behind this result, note that the owner pays the manager a compensation premium only if she approves the project. Therefore, the owner can reduce her expected compensation cost by reducing the set of states for which she approves the project. Such a strategy is costly, however, because the owner forgoes some positive net present value investments. The owner's choice of hurdle rate, identified in Observation 1, balances her conflicting objectives of maximizing the investment returns, while minimizing the expected compensation payment to the manager.

To provide more intuition behind Expression (6), I evaluate how the owner's choice of hurdle rate affects her expected payoff. First, if the owner increases the hurdle rate r^{o} by a small amount ε, her expected payoff decreases by $k \cdot (r^{o} - c) \cdot f(r^{o}) \cdot \varepsilon$ because of the expected net present value forgone. At the same time, the higher hurdle rate benefits the owner because she can reduce the manager's compensation by $v \cdot k \cdot \varepsilon$ whenever the project's rate of return r is in the investment region (i.e., $r > r^{o}$). Therefore, if the owner increases the hurdle rate by ε, then her expected compensation cost decreases by $v \cdot k \cdot \varepsilon \cdot [1 - F(r^{o})]$. At the optimum, the owner sets the marginal cost of underinvestment $k \cdot (r^{o} - c) \cdot f(r^{o})$ equal to its marginal benefit $v \cdot k \cdot [1 - F(r^{o})]$.

Optimal Investment in the Presence of a Managerial Retention Problem

With the above benchmark in place, I derive the owner's project approval rule when the manager can profitably pursue the project in an outside venture (i.e., $\delta > 0$), and therefore the owner faces a managerial retention problem. In solving her optimization problem, the owner must now take into account the manager's retention constraint. The technical challenge in this setting is that the retention constraint makes the manager's reservation utility dependent on his private information.[14] As a consequence, the standard methodology for solving the adverse-selection problems no longer applies.

It turns out, however, that the owner's optimal investment policy is again upper-tailed—that is, the owner approves the project only if the reported rate of return exceeds a hurdle rate. Recall that r^{o} denotes the hurdle rate that the owner requires for approving the project in the absence of a managerial retention problem (i.e., when $\delta = 0$). Let r^{*} denote the owner's choice of hurdle rate when the she faces a managerial retention threat (i.e., when $\delta > 0$). The result below shows that the owner's optimal choice of hurdle rate depends on the severity of the managerial retention problem, as reflected by the magnitude of δ.

Proposition 1: (i) When the managerial retention problem is not too severe (i.e., δ is less than a threshold level δ_c), the owner's optimal choice of hurdle rate is the same as when $\delta = 0$. That is, $r^{*} = r^{o}$ when $\delta \leq \delta_c$.

---

[14] Lewis and Sappington (1989) analyze a regulation setting in which a regulated firm's reservation price varies with its type.

(ii) When the managerial retention problem is severe (i.e., when δ exceeds δ_c), the owner's optimal choice of hurdle rate solves the following equation:

$$r^* = c + v \cdot \frac{F(\rho) - F(r^*)}{f(r^*)}, \tag{7}$$

where:

$$\rho \equiv \frac{(c - v \cdot r^*)}{(\delta - v)}.^{15} \tag{8}$$

Figure 1 clarifies the intuition behind Part (i) of Proposition 1. The solid line in Figure 1 displays the manager's net payoff (i.e., his compensation less his cost of effort) from an incentive plan that induces him to communicate the project's rate of return truthfully. Note that when the owner approves the project (i.e., when $r > r^*$), the manager's net payoff $v \cdot k \cdot (r - r^*)$ is increasing at the rate of $v \cdot k$ in his private information r. As discussed earlier, this is a necessary condition for the manager to report his private information truthfully. That is why I refer to the manager's payoff depicted by the solid line in Figure 1 as the manager's *incentive premium*. The higher the owner sets the capital investment hurdle rate r^*, the less incentive premium she has to pay to the manager.

The dotted line in Figure 1 shows the minimum net payoff that the manager requires to stay with the firm, $\delta \cdot k \cdot (r - c/\delta)$. If the project's outside rate of return $\delta \cdot r$ is less than the manager's cost of capital c, the manager has no incentive to take the project outside. Thus, the owner needs to consider the manager's retention constraint only when r exceeds c/δ, the manager's break-even point.

The key point in Figure 1 is that when the project is relatively less profitable outside the firm (i.e., when δ is small), the incentive premium corresponding to the hurdle rate r^o (shown in the solid line) exceeds the minimum net payoff necessary to retain the manager (shown in the dotted line). That is:

$$v \cdot k \cdot (r - r^o) > \delta \cdot k \cdot (r - c/\delta) \quad \text{for each r.} \tag{9}$$

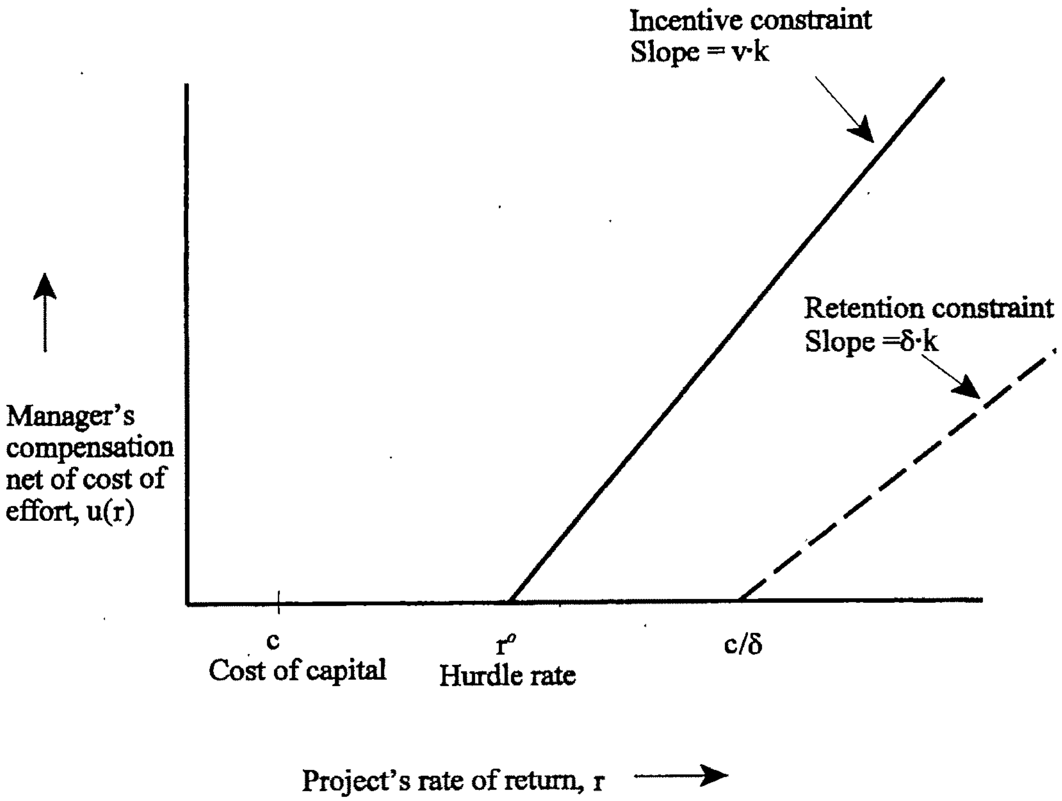
Consequently, as illustrated in Figure 1, the hurdle rate identified in Observation 1 remains optimal for small values of δ.

When the severity of the managerial retention problem (as reflected by the magnitude of δ) reaches a certain threshold level, however, Inequality (9) fails to hold because the optimal incentive plan for the $\delta = 0$ case does not provide enough compensation to meet the manager's retention constraint. To deter the manager from departing with the project, the owner must provide him with additional compensation, which I refer to as the manager's *retention premium*. This case is illustrated in Figure 2, where, as before, the solid line displays the net payoff necessary to induce the manager to report truthfully, and the dotted line represents the net payoff necessary to retain him.

---

[15] There are two possible cases, depending on whether the manager's marginal cost of effort v is less or greater than c/r^o. In the remainder of the text, I assume that $v \leq c/r^o$. The solution to the case $v > c/r^o$ appears in the Appendix. However, as shown in the Appendix, all of the subsequent results (i.e., Propositions 2 and 3) hold for both cases.

FIGURE 1
Optimal Investment and Incentive Plans
(Mild Retention Problem)



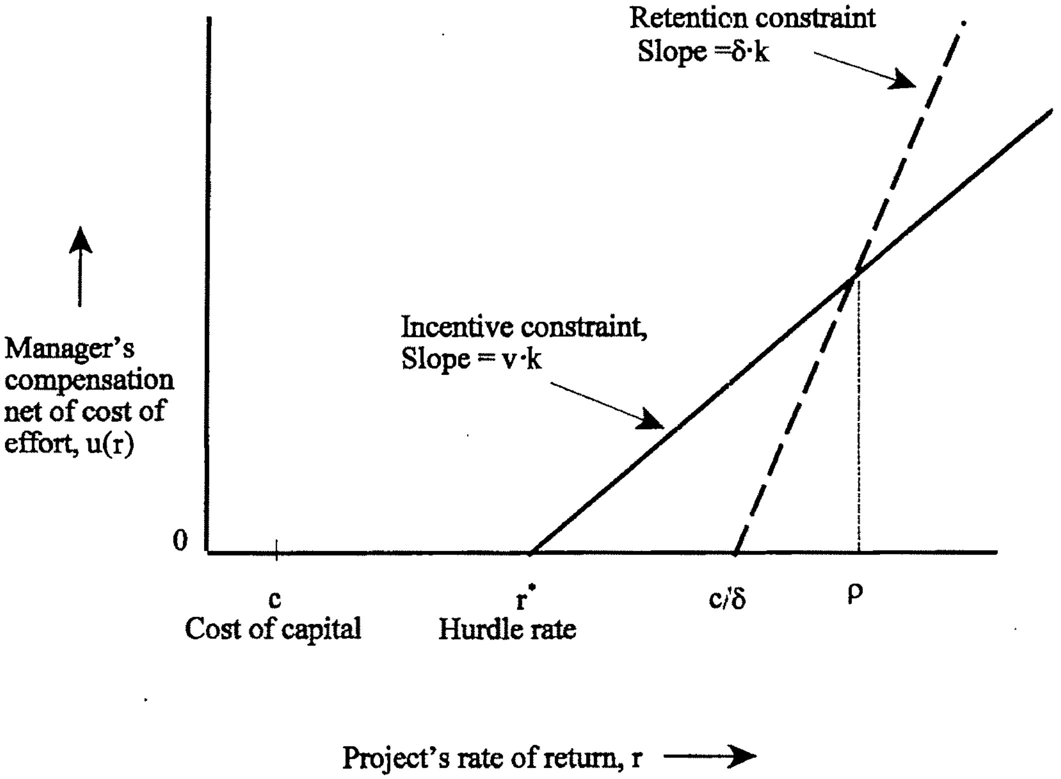
To interpret the expression for the owner's optimal choice of hurdle rate (i.e., Expression (7)), I consider how the owner's expected payoff varies with her choice of hurdle rate. If she increases the hurdle rate r^* by a small amount ε, the owner's expected profit decreases by $k \cdot (r^* - c) \cdot f(r^*) \cdot \varepsilon$ because of the positive net present value forgone. Since the manager earns an incentive premium only when the owner approves the project, the increased hurdle rate also benefits the owner because she pays a lower incentive premium to the manager.

However, the owner's choice of hurdle rate no longer determines the manager's compensation for high values of r. As shown in Figure 2, when the project's rate of return r exceeds ρ, the manager's incentive premium is less than the minimum payment necessary to retain him. Consequently, when the project's rate of return r exceeds ρ, the retention constraint, rather than the incentive compatibility constraint, determines the manager's compensation. As a result, the owner's choice of hurdle rate r^* affects the manager's compensation only when the project's rate of return r is between r^* and ρ.[16] Therefore, when the

---

[16] Of course, as reflected in Expression (8), the rate of return beyond which the retention constraint determines the manager's compensation (i.e., the point ρ in Figure 2) is itself a function of the hurdle rate r^*.

FIGURE 2
Optimal Investment and Incentive Plans (Severe Retention Problem)



owner increases the hurdle rate by a small amount ε, her expected compensation cost decreases by $v \cdot k \cdot [F(\rho) - F(r^*)] \cdot \varepsilon$. At the optimum, Expression (7) equates the marginal cost of underinvestment $k \cdot (r^* - c) \cdot f(r^*)$ to its marginal benefit $v \cdot k \cdot [F(\rho) - F(r^*)]$.

Comparing the owner's choice of hurdle rate in Expression (7) with that in Expression (6) reveals that $r^* \leq r^o$—that is, the owner uses a lower hurdle rate for project approval when the retention problem is severe. In fact, I obtain the following monotonicity result:

Proposition 2: The optimal hurdle rate decreases in the severity of the managerial retention problem (i.e., r^* decreases in δ).

Proposition 2 establishes that as the project becomes relatively more valuable outside the firm, and hence the managerial retention problem becomes more severe, the owner invests more efficiently (i.e., she uses a hurdle rate closer to her true cost of capital). At first glance, this result may seem counterintuitive. However, the point to note is that the owner sets a hurdle rate higher than her cost of capital to reduce the managerial compensation cost. As the project becomes relatively more profitable outside the firm, the manager demands and receives higher compensation to remain with the firm. Since the owner's choice of hurdle rate has no direct effect on the manager's retention premium, the component of the manager's compensation that the owner can control through her choice of hurdle rate decreases (i.e., the incentive premium *in excess of* the retention premium decreases). Put another way, while the marginal cost of underinvestment remains unchanged, the marginal benefit of underinvestment decreases as the severity of the managerial retention problem, δ, increases. Consequently, the owner finds it optimal to introduce less underinvestment.

This result provides a testable prediction regarding the cross-sectional variation in firms' choices of hurdle rates for approving capital investment projects. In particular, it predicts that, all else being equal, firms will use lower hurdle rates for internal capital investment decisions when their investment opportunities (1) require specific managerial expertise, and (2) can easily be exploited in outside ventures.

IV. DELEGATED INVESTMENT DECISIONS

In the previous section, I considered a centralized capital budgeting setting in which the manager reports his private information to the owner, who then decides whether to approve the project. This arrangement represents the typical capital budgeting process within a centralized firm in which the headquarters makes the capital investment decisions based on the divisional managers' proposals.

In this section, I investigate a decentralized organization in which the owner delegates the investment decision to the manager, who does *not* communicate his private information about the project to the owner. We know from the revelation principle that the owner's expected payoff from a decentralized organization can never exceed her expected payoff from a centralized organization. I examine whether the owner can attain the same expected payoff with delegation as she would earn under centralized capital budgeting described in the previous section.

It is important to analyze the delegated setting because shareholders usually assign investment decision rights to corporate executives. Furthermore, in many multidivisional firms, corporate headquarters delegate investment decisions to divisional managers. To align managers' incentives with shareholders' objectives, such firms compensate their managers on the basis of performance measures such as share price, accounting income, residual income, and return on assets.

Under delegation, the owner hands over the investment decision to the manager. To ensure that the manager makes the desirable investment decision and exerts a high level of effort, the owner designs a managerial compensation package (i.e., a performance measure and compensation rule). The residual income performance measure—that is, operating income less an interest charge for the use of capital—has recently received considerable attention in both the academic and practitioner literatures.[17] Relative to operating income, which does not deduct any charge for the use of capital, the residual income performance measure has the advantage that shareholders can tailor the capital charge so that the manager internalizes their investment objectives. To investigate the incentive properties of the residual income performance measure in my model, I consider managerial compensation contracts of the form:

$$s = \alpha + \beta \cdot RI. \tag{10}$$

In Expression (10), α denotes the manager's fixed salary, $\beta \cdot RI$ is a bonus payment based on residual income RI, and β is the manager's bonus parameter.

---

[17] For analytical models on the incentive properties of residual income, see Rogerson (1997), Reichelstein (1997), Christensen et al. (2002), and Dutta and Reichelstein (2002). Ittner and Larcker (1998) provide a comprehensive review of different variants of residual income-type performance measures used in practice. For instance, Stewart (1991) recommends a particular form of residual income, namely economic value added (or EVA).

In my one-period model, operating income equals net cash flow, that is:

$$\begin{aligned} \mathrm{Inc} &= x - k \cdot I \\ &= a + r \cdot k \cdot I. \end{aligned}$$

Therefore, residual income corresponding to the *capital charge rate* of h is:

$$\mathrm{RI} = a + (r - h) \cdot k \cdot I. \tag{11}$$

In the delegated investment setting, I call a managerial compensation contract *optimal* if it (1) satisfies the manager's participation and retention constraints, and (2) generates the same expected payoff for the owner as the optimal incentive plan under centralized capital budgeting.

To investigate the owner's choice of the optimal managerial compensation contract, consider first the benchmark problem in which the project has no value outside the firm (i.e., $\delta = 0$). Given the linear compensation contract in Equation (10) and the residual income performance measure in Equation (11), the manager will invest if and only if the project's rate of return r exceeds h, the rate at which the owner charges the manager for the use of capital. If the owner uses operating income as a managerial performance measure (i.e., if she sets h equal to zero), the manager will accept the project whenever $r > 0$, which is clearly suboptimal. The manager overinvests because he does not internalize the owner's cost of capital.

On the other hand, if the capital charge rate h equals the owner's cost of capital c, the manager will undertake the project whenever its net present value is positive, and reject the project when it is negative. In the absence of an agency problem (i.e., both the owner and manager know the project's rate of return), this investment decision rule is clearly efficient. In my model, however, the manager has private information and consequently, as shown in the previous section, the owner finds it optimal to invest less than the efficient level.

If the owner sets the capital charge rate equal to the hurdle rate r^o, the manager will adopt the optimal project acceptance decision described in the previous section. Furthermore, the manager will exert the maximum level of effort if the manager's bonus parameter β equals his marginal cost of effort v. If the owner sets the fixed salary α equal to the manager's reservation wage, the linear compensation contract in Equation (10) will provide the manager with a compensation of $v \cdot \bar{a} + \max\{0, v \cdot k \cdot (r - r^o)\}$, which is the same as his compensation from the optimal incentive plan under centralized capital budgeting. This leads to the following result:

> **Observation 2:** When the project has no value outside the firm (i.e., when $\delta = 0$), a linear compensation contract based on residual income is optimal.

I now investigate whether a bonus contract of the form in Equation (10) can be optimal when the owner faces a managerial retention problem (i.e., when $\delta > 0$). Recall from Proposition 1 that under centralized capital budgeting, the owner's choice of hurdle rate depends on the severity of the managerial retention problem (as captured by the magnitude of δ). When the retention problem is mild (specifically, when $\delta \leq \delta_c$), Proposition 1 shows that the owner sets the hurdle rate as though $\delta = 0$ (i.e., as though the retention constraint does not apply). Consequently, it follows from the discussion preceding Observation 2 that when the retention problem is mild, a linear contract based on residual income can give the manager optimal effort and investment incentives.[18]

When the managerial retention problem is severe (specifically, when $\delta > \delta_c$), the manager's compensation under centralized capital budgeting is determined by (1) the incentive constraint for low values of r, and (2) the retention constraint for high values of r. Consequently, as shown in Figure 2, the manager's compensation under centralized capital budgeting is a convex function of his private information r. This suggests that a simple linear residual-income-based bonus contract may no longer be optimal.

To see that this is indeed the case, note that the manager will exert the optimal amount of effort only if his bonus parameter β is (weakly) greater than his marginal cost of effort v. At the same time, the manager will adopt the optimal project acceptance rule only if the capital charge rate equals r^*. Furthermore, the manager's compensation contract must be such that his net payoff from accepting the contract is at least equal to what he can earn by quitting and pursuing the project on his own. A linear compensation contract of the form in Equation (10) cannot satisfy these incentive and retention requirements and, at the same time, ensure that the owner earns the same expected payoff as she would have earned under centralized capital budgeting.

As shown in Figure 2, the manager's payoff under centralized capital budgeting is a piecewise linear and convex function of the project's rate of return. This suggests that to attain the same payoff as she could have earned under centralized capital budgeting, the owner must supplement the linear compensation contract in Equation (10) with an additional bonus component that kicks in after the manager's performance has reached a certain threshold. The result below shows that the owner can achieve this objective by including a properly calibrated option component in the manager's compensation contract.

Proposition 3: (i) A linear contract of the form in Equation (10) is optimal if and only if the managerial retention problem is mild (i.e., $\delta \leq \delta_c$).

(ii) When the managerial retention problem is severe (i.e., $\delta > \delta_c$), a compensation contract of the following form is optimal:

$$s = \alpha + \beta \cdot [x - (1 + h) \cdot k \cdot I] + \gamma \cdot \max\{0, x - \psi\}. \qquad (12)$$

I interpret the third term on the right-hand side of Expression (12) as the manager receiving γ options, each with an exercise price of ψ. Since the manager will exercise options only when the end-of-period operating cash flow x exceeds the option exercise price ψ, this additional bonus component kicks in only after the manager's performance has reached a certain threshold. To ensure that the manager exerts a high level of effort and makes the optimal investment decision, the owner sets the bonus parameter β equal to the manager's marginal cost of effort v, and the capital charge rate h equal to the hurdle rate r^*. The owner can then choose the coefficient γ and the option exercise price ψ so that the manager's compensation from the compensation contract in Equation (12) is identical to that from the optimal incentive plan under centralized capital budgeting. These choices of the compensation parameters ensure that the managerial compensation contract in Equation (12) generates the same expected profit for the owner as the optimal incentive plan described in the previous section.[19]

---

[18] If the managerial retention problem is mild, a linear contract is optimal in my setting. In contrast, some of the earlier studies investigate compensation contracts that are optimal *within* the class of linear contracts (Bushman and Indjejikian 1993; Feltham and Xie 1994; Indjejikian and Nanda 1999; Datar et al. 2001; Christensen et al. 2002; Dutta and Reichelstein 2002).

Although the optimal values of the fixed salary α and the bonus parameter β are independent of the severity of the managerial retention problem δ, the optimal value of the option coefficient γ is increasing in δ. Thus, the relative weight on the option component increases as the retention problem becomes more severe. This prediction is consistent with the common observation that new economy, high-tech firms are increasingly compensating their managers with stock options, because their investment opportunities (1) depend on managerial expertise, and (2) can easily be exploited in outside ventures. Core and Guay (2001) provide empirical evidence supporting this prediction.

In my model, the owner includes an option component in the manager's compensation contract not so much to generate desirable effort and investment incentives as to retain valuable managerial human capital and the associated investment projects. This is consistent with the empirical results of Ittner et al. (2001), who find that option grants increase managerial retention and attraction. Many in the business press claim that the primary purpose of stock-options-based compensation plans is to attract or retain managers, rather than to motivate them (e.g., Battey 2000; Portnot and Moltzen 2000). As Ittner et al. (2001) observe, however, if the firm's objective is merely to retain or attract managers, it is the *level* and not the *form* of compensation that should matter. It has been puzzling, then, why firms would use stock options rather than fixed salaries to attract and retain managers. My model provides a theoretical rationale for why stock options help attract and retain managers.

My one-period analysis assumes that the manager is irreplaceable because he has some unique expertise essential to execute the firm's investment project. In multiperiod settings, the manager is likely to become increasingly irreplaceable as he develops more firm-specific expertise over time. This suggests that the managerial retention problem may increase over time. My results provide a testable prediction that the relative weight on the option component in the manager's compensation contract will increase with managerial tenure.

V. CONCLUSION

I have used an agency model to examine how managerial retention and incentive problems affect a firm's choice of (1) hurdle rate for capital investment when the owner makes the investment decision, and (2) management compensation plan when the owner delegates the investment decision to the manager. When the owner retains the investment authority, my analysis shows that as the managerial retention problem becomes more severe (i.e., as the manager finds it easier to leave the firm and pursue the project in outside ventures), the owner requires a lower hurdle rate for project approval.

When the owner delegates the investment decision to the manager, I show that if the managerial retention problem is not too severe, a residual-income-based bonus contract can provide the manager with optimal effort and investment incentives. In contrast, when the managerial retention threat becomes more serious, the owner must include an option-like component in the manager's compensation contract in order to retain the manager and related investment opportunities, and to provide him with appropriate effort and investment incentives. This theoretical result is consistent with the observation that new-economy, high-tech firms are increasingly compensating their managers with stock options, because the investment opportunities of these firms depend on managerial expertise, and can easily be exploited in outside ventures.

---

[19] Since the manager's cost of effort is linear, my model amounts to a binary effort choice problem in which the owner seeks to induce the manager to exert high effort. This aspect of the model (and not the linearity assumption, per se) simplifies my analysis. If the manager's disutility for effort was nonlinear, then all of my results would remain unchanged as long as the marginal return of managerial effort is sufficiently large so that the owner induces the manager to exert the maximum level of effort. In general, however, the owner's optimal choice of the managerial effort will depend on the reported r. Although the analysis becomes considerably more complicated, the result that the optimal hurdle rate decreases in the retention risk continues to hold. Furthermore, an appropriately designed option-type contract of the form in Equation (12) continues to dominate *any* linear compensation contract.

My analysis readily extends to a two-period setting in which the manager learns about a second investment opportunity at the end of the first period. If the manager can again walk away with the second investment opportunity, then the manager's optimal compensation contract for the second period will remain qualitatively unchanged from the one derived in my simpler one-period model. An initial investment in many instances is required to identify future investment opportunities. The real option literature characterizes such initial investments as embedded with *growth options*. Future research could extend my framework to investigate optimal (1) capital budgeting for projects that come embedded with future growth options, and (2) management compensation plans to address related managerial incentive and retention problems.

I assume that the investment project has a useful life of only one period. Future research could extend this analysis to a more realistic setting in which the investment payoffs arrive over a number of periods. In a multiperiod extension of my model, one can examine the incentive properties of different accounting rules (e.g., depreciation methods). The recent multiperiod literature that investigates the incentive properties of different depreciation rules has not yet considered the effect of managerial retention threats.

APPENDIX

Proof of Observation 1

Following standard arguments from the adverse-selection literature (e.g., Baron and Myerson 1982), I can show that, when the induced effort choice is interior (i.e., $a(r) \in (0, \bar{a})$), a direct revelation mechanism is incentive compatible if and only if (i) $u'(r) = v \cdot k \cdot I(r)$, and (ii) $I(r)$ is (weakly) increasing. For the two boundary efforts $\{0, \bar{a}\}$, these conditions are sufficient, but not necessary. When the induced effort choice is $\bar{a}$, the manager cannot overstate his type, and so only the downward incentive constraints apply. The incentive compatibility constraint is equivalent to Condition (ii) above and $u'(r) \geq v \cdot k \cdot I(r)$. When the induced effort choice is zero, the manager cannot understate his type, and the condition $0 \leq u'(r) \leq v \cdot k \cdot I(r)$ is equivalent to Constraint (IC).

Combining the above results, it follows that Constraint (IC) is equivalent to the monotonicity constraint that I(r) is increasing, and the condition that:

$$u'(r) = v \cdot k \cdot I(r) \cdot [J(a(r)) + \Delta(r)] \quad \text{for all } r, \tag{A1}$$

where the indicator function $J(a) \in [0, 1]$ must equal 1 whenever $a > 0$, and the nonnegative valued function $\Delta(\cdot)$ must equal zero whenever $a < \bar{a}$. Integrating Equation (A1) yields:

$$u(r) = u(\underline{r}) + v \cdot k \cdot \int_{\underline{r}}^{r} I(z) \cdot [J(a(z)) + \Delta(z)]\, dz \quad \text{for all } r. \tag{A2}$$

It follows from Equation (A2) that the participation constraint will hold for each type if it is satisfied for the lowest type. Therefore, I can drop Constraint (IR) by setting $u(\underline{r}) \geq 0$. To minimize the manager's expected compensation, the owner will optimally choose

$u(\underline{r}) = 0$ and set $\Delta(r) = 0$ for each r. Substituting these choices in Equation (A1), using the definition $u(r) \equiv s(r) - v \cdot a(r)$ and integrating by parts yields:

$$\int_{\underline{r}}^{\bar{r}} s(r)\, f(r)\, dr = \int_{\underline{r}}^{\bar{r}} [v \cdot a(r) + v \cdot k \cdot I(r) \cdot J(a(r)) \cdot H(r)]\, f(r)\, dr.$$

As a result, the owner's optimization problem simplifies to:

$$\max_{a(r), I(r)} \int_{\underline{r}}^{\bar{r}} (a(r) \cdot (1 - v) + I(r) \cdot k \cdot [r - c - v \cdot J(r) \cdot H(r)])\, f(r)\, dr.$$

I first solve the above program without the constraint that I(r) is increasing, and then verify that the optimal solution to the relaxed program satisfies this constraint.

Consider first the choice of the optimal managerial effort. If the following condition:

$$\bar{a} \cdot (1 - v) \geq v \cdot k \cdot H(c) \tag{A3}$$

holds, then $a(r) = \bar{a}$ is the optimal choice of the managerial effort for each r. Recall that $J(a(r)) \leq 1$, and hence $a(r) = \bar{a}$ is optimal if $\bar{a} \cdot (1 - v) \geq v \cdot k \cdot H(r) \cdot I(r)$. Clearly, this condition holds for all $r \leq c$ because the owner will optimally set $I(r) = 0$ in this region. It also holds for all $r \geq c$ and each $I(r) \in \{0,1\}$, because (i) it holds at $r = c$ from (A3), and (ii) the inverse hazard rate H(r) is decreasing.

Since it is optimal to induce $a > 0$, $J(a(r)) = 1$. Substituting this in the owner's objective function reveals that the owner will optimally approve the project if and only if:

$$r - c - v \cdot H(r) \geq 0.$$

For brevity, let G(r) denote the left-hand side of the above inequality. Note that (i) $G(c) < 0$, (ii) $G((\bar{r}) > 0$ since $H(\bar{r}) = 0$, and (iii) $G'(r) > 0$ since $H'(r) < 0$. These facts imply that there exists a unique $r^o \in (c,\bar{r})$ such that the owner will approve the project if and only if $r \geq r^o$, thereby proving that I(r) is increasing. This completes the proof.

Proof of Proposition 1

It is again optimal for the owner to induce the maximum amount of effort from the manager for each r provided that (A3) holds. Therefore, the incentive compatibility constraint is equivalent to the condition that I(r) is increasing and:

$$u(r) = v \cdot k \cdot \int_{\underline{r}}^{r} I(z) \cdot [1 + \Delta(z)]\, dz \tag{A4}$$

for any function $\Delta(\cdot) \geq 0$. Since I(r) must be increasing, Condition (A4) implies that for some r^*:

$$u(r) = \begin{cases} k \cdot [v \cdot (r - r^*) + D(r)] & \text{if } r \geq r^* \\ 0 & \text{otherwise} \end{cases}$$

where:

$$D(r) \equiv v \cdot \int_{r^*}^{r} \Delta(z)\, dz \geq 0.$$

Note that $D'(r) = v \cdot \Delta(r) \geq 0$. Consequently, the owner's optimization problem becomes:

$$\max_{r^*, D(r)} k \cdot \int_{r^*}^{\bar{r}} [r - c - v \cdot (r - r^*) - D(r)]\, f(r)\, dr$$

subject to:

$$\begin{array}{rll} k \cdot [v \cdot (r - r^*) + D(r)] \geq k \cdot (\delta \cdot r - c) & \text{for all } r \geq r^* & \text{(RC)} \\ D(r) \geq 0 & \text{for all } r & \text{(i)} \\ D'(r) \geq 0 & \text{for all } r. & \text{(ii)} \end{array}$$

From Observation 1, we know that $r^* = r^o$ will be the optimal solution whenever $D(r) = 0$ for each r satisfies Constraint (RC). To proceed with the rest of the proof, consider the following two cases:

Case I: $c - v \cdot r^o \geq 0$.

Part (i): If $\delta \leq \delta_c$, then $r^* = r^o$ is the optimal solution, where $\delta_c \equiv v + \frac{c - v \cdot r^o}{\bar{r}}$.

Proof: It suffices to show that $D(r) = 0$ satisfies Constraint (RC) for each r. Therefore, I need to show that the inequality:

$$(v - \delta) \cdot r + (c - v \cdot r^o) \geq 0 \qquad \text{(A5)}$$

holds for each r above the manager's break-even point $c \cdot \delta^{-1}$. Clearly, Inequality (A5) holds for all $\delta \leq v$. When $\delta > v$, it will hold for each type if it holds for the highest type $\bar{r}$ and for $\delta = \delta_c$. Inequality (A5) holds for the highest type because $(v - \delta_c) \cdot \bar{r} + (c - v \cdot r^o) = 0$ by the definition of δ_c. This proves Part (i) of Proposition 1.

Part (ii): If $\delta \geq \delta_c$, then the optimal hurdle rate is given by Expression (7) in Proposition 1.

Proof: I first show that the optimal hurdle rate is always below the manager's break-even point $c \cdot \delta^{-1}$. To prove this by contradiction, suppose $r^* > c \cdot \delta^{-1}$. Constraint (RC) then requires that $D(r) = (\delta \cdot r - c) - v \cdot (r - r^*) > 0$ for each $r \geq r^*$. Substituting this in the owner's objective function leads to the following expression for her expected profits:

$$\pi(r^*) \equiv \int_{r^*}^{\bar{r}} r \cdot (1 - \delta)\, f(r)\, dr.$$

Clearly, $\pi(\cdot)$ is decreasing, which contradicts the hypothesis that $r^* > c \cdot \delta^{-1}$ is optimal.

For each $r^* \in [c, c \cdot \delta^{-1}]$, let $\rho \geq r^*$ denote the highest type for which $D(r) = 0$ satisfies constraint (RC). That is:

$$v \cdot (\rho - r^*) = \delta \cdot \rho - c. \tag{A6}$$

The owner will optimally set $D(r) = 0$ for all $r < \rho$, and $D(r) = (\delta - v) \cdot k \cdot (r - \rho)$ for all $r \geq \rho$. This choice of $D(r)$ satisfies the monotonicity Constraint (ii) in the owner's optimization problem. The optimal hurdle rate solves the following optimization problem:

$$\max_{r^*} \pi(r^*) \equiv \int_{r^*}^{\rho} [r - c - v \cdot (r - r^*)]\, f(r)\, dr + \int_{\rho}^{\bar{r}} r \cdot (1 - \delta)\, f(r)\, dr.$$

Differentiating yields:

$$\pi'(r^*) = f(r^*) \cdot \left(v \cdot \frac{F(\rho) - F(r^*)}{f(r^*)} - (r^* - c) \right).$$

First, note that $\pi'(c) > 0$. Second, $[F(\rho) - F(r)] \cdot [f(r)]^{-1} \geq H(r) \equiv [1 - F(r)] \cdot [f(r)]^{-1}$, which implies that $\pi'(r^o) < 0$. Furthermore, $[F(\rho) - F(r^*)] \cdot [f(r^*)]^{-1}$ is decreasing in r^* if the inverse hazard rate $H(\cdot)$ is decreasing. It thus follows that the term inside the outer parenthesis is decreasing in r^*, and hence the equation $\pi'(r^*) = 0$ has a unique solution in (c, r^o). This proves Part (ii) of Proposition 1.

Case II: $c - v \cdot r^o < 0$.

The owner's optimal investment policy depends on whether $\delta < c/r^o$, $\delta \in [c/r^o, v]$, or $\delta > v$. It is easy to verify that (i) $r^* = r^o$ is the optimal solution when $\delta < c/r^o$, and (ii) the optimal solution is the same as Part (ii) of Proposition 1 when $\delta > v$.

Consider now the case $\delta \in [c/r^o, v]$. I show that the optimal hurdle rate always exceeds the manager's break-even point $c \cdot \delta^{-1}$. To prove this by contradiction, suppose $r^* < c \cdot \delta^{-1}$. Then $D(r) = 0$ will satisfy Constraint (RC) for each r. Therefore, the owner's expected profit as a function of r^* is increasing for all $r^* < r^o$. Since $c \cdot \delta^{-1} < r^o$, it follows that the owner's expected profit is increasing for all values of r^* below $c \cdot \delta^{-1}$, which contradicts the hypothesis that $r^* < c \cdot \delta^{-1}$ is optimal.

For each $r^* \geq c \cdot \delta^{-1}$, define σ by the equation:

$$v \cdot (r^* - \sigma) = \delta \cdot r^* - c.$$

For a given r^*, the owner will optimally set $D(r) = v \cdot (r^* - \sigma)$ for all $r \geq r^*$, and $D(r) = 0$ for all other values of r. Note that manager types in the interval $[c \cdot \delta^{-1}, r^o)$ will decide to leave. The owner will therefore choose $r^* \geq c \cdot \delta^{-1}$ to maximize:

$$\pi(r^*) \equiv \int_{r^*}^{\bar{r}} [r - c - v \cdot (r - \sigma(r^*))]\, f(r)\, dr.$$

Differentiating and applying Leibitz's rule yields:

$$\pi'(r^*) = f(r^*) \cdot [(v - \delta) \cdot H(r^*) - r^* \cdot (1 - \delta)]. \tag{A7}$$

Note that $(v - \delta) \cdot H(r^o) - (1 - \delta) \cdot r^o < v \cdot H(r^o) - r^o + \delta \cdot r^o < v \cdot H(r^o) - r^o + c = 0$, and hence the right-hand-side of Expression (A7) is negative when evaluated at $r^* = r^o$. Furthermore, the term inside the square brackets is strictly decreasing in r^*. These two facts imply that the equation $\pi'(r^*) = 0$ has at most one solution in the interval $[c \cdot \delta^{-1}, r^o)$. The optimal hurdle rate is either equal to $c \cdot \delta^{-1}$, or is given by the unique solution of the equation $\pi'(r^*) = 0$.

Proof of Proposition 2

To facilitate the presentation of the proof, I combine Cases I and II in the proof of Proposition 1 by defining:

$$\delta_L \equiv \min\left\{\frac{c}{r^o}, v + \frac{c - v \cdot r^o}{\bar{r}}\right\}$$

$$\delta_H \equiv \max\left\{v, v + \frac{c - v \cdot r^o}{\bar{r}}\right\}.$$

When $\delta \geq \delta_L$, the optimal hurdle rate is independent of δ because it is equal to r^o. If $\delta \geq \delta_H$, then the optimal hurdle rate solves the equation:

$$v \cdot \hat{H}(r^*(\delta),\delta) - r^* + c = 0, \tag{A8}$$

where $\hat{H}(r^*(\delta),\delta) \equiv [F(\rho(\delta)) - F(r^*)]/f(r^*(\delta))$ and ρ is given by Expression (A6). Differentiating Equation (A8) yields:

$$v \cdot \left(\frac{\partial \hat{H}}{\partial r^*} \cdot \frac{\partial r^*}{\partial \delta} + \frac{\partial \hat{H}}{\partial \delta}\right) - \frac{\partial r^*}{\partial \delta} = 0.$$

It is easy to verify that $\hat{H}(r^*,\delta)$ is decreasing in r^* if the inverse hazard rate $H(\cdot)$ is decreasing. Furthermore, $\hat{H}(r^*,\delta)$ is also decreasing in δ because $\rho(\delta)$ is decreasing. These two facts imply that the optimal hurdle rate r^* is decreasing in δ for $\delta \geq \delta_H$.

In the region $\delta \in [\delta_L, \delta_H]$, the optimal hurdle rate either (i) is equal to $c \cdot \delta^{-1}$, or (ii) solves the equation $\pi'(r^*) = 0$, where $\pi'(\cdot)$ is given by Expression (A7). Clearly, the optimal hurdle is decreasing in δ in case (i). In case (ii), the derivative of the optimal hurdle rate (with respect to δ) has the same sign as the sign of $r^* - H(r^*)$, which is negative because:

$$r^* = (v - \delta) \cdot (1 - \delta)^{-1} \cdot H(r^*) < H(r^*).$$

Proof of Observation 2

If the owner chooses the bonus coefficient $\beta = v$, the manager will choose $a = \bar{a}$. A necessary and sufficient condition to ensure that the manager has optimal investment incentives is that the manager is indifferent between accepting and rejecting the marginal project (the project that returns r^o). Setting the capital charge rate h equal to the hurdle rate r^o ensures that the manager invests optimally. The fixed salary $\alpha = 0$ guarantees that the manager's participation constraint is satisfied. Since each type of manager earns the same compensation premium as he earns in the optimal incentive plan under centralization, the owner's expected profit from the above contract is identical to that from the optimal incentive plan under centralization.

Proof of Proposition 3

Step I: A residual-income-based linear contract generates optimal effort and investment incentives for the manager if and only if $\delta \leq \delta_H$.

Proof: I first prove the "if" part. Recall from the proof of Proposition 2 that when $\delta \leq \delta_L$, the optimal incentive plan under centralization is the same as the one in the absence of the retention constraint. Therefore, the linear contract in Observation 2 will continue to generate optimal incentives. When $\delta \in [\delta_L, \delta_H]$, the optimal hurdle rate is either (i) equal to $c \cdot \delta^{-1}$, or (ii) given by the solution to the equation $r^* = (v - \delta) \cdot (1 - \delta)^{-1} \cdot H(r^*)$. In case (i), a linear contract with $\alpha = 0$, $\beta = v$, and $h = r^*$ satisfies the manager's participation and retention constraints, and provides him with optimal effort and investment incentives. Since the manager's compensation is identical to that in the optimal incentive plan under centralization, the above contract is optimal. In case (ii) a linear contract with $\alpha = 0$, $\beta = v$, and $h = \sigma$, where σ was defined in the proof of Proposition 2, satisfies the manager's retention constraint for all $r \geq r^*$. The delegated contract is again optimal because it induces the optimal effort and investment incentives, and gives the manager the same compensation as the optimal incentive plan under centralization.

To prove the "only if" part, I will show that a linear incentive contract cannot be optimal when $\delta > \delta_H$. To provide optimal investment incentives, the capital charge rate must equal the hurdle rate r^*. This implies that the manager will earn rent $\beta \cdot k \cdot (r - r^*)$, and hence the retention constraint requires that $\beta \cdot k \cdot (r - r^*) \geq k \cdot (\delta \cdot r - c)$ for all r. Recall that the optimal incentive plan under centralization provides a compensation premium of $v \cdot k \cdot (r - r^*)$ for all $r < \rho$ and $k \cdot (\delta \cdot r - c)$ for all $r \geq \rho$. It is impossible to find a coefficient β that satisfies the retention constraint and at the same time ensures that the manager's expected compensation is the same as in the optimal incentive plan under centralization.

Step II

For $\delta > \delta_H$, I claim that the contract of the form in Equation (12) is optimal when the fixed payment $\alpha = 0$, the bonus coefficient $\beta = v$, the number of options $\gamma = \delta - v$, the option exercise price $\psi = \bar{a} + (1 + \rho) \cdot k$. and the capital charge rate $h = r^*$. The above choice of capital charge rate ensures that the manager will make the optimal investment decision. The choice of bonus parameter guarantees that the manager will contribute $a = \bar{a}$. Since the manager will exercise the options only when $r \geq \rho$, he will earn $v \cdot k \cdot (r - r^*)$ for all $r < \rho$ and $k \cdot (\delta \cdot r - c)$ for all $r \geq \rho$. Thus, the compensation contract in Equation (12) is optimal because it: (i) provides the manager with the same amount of compensation as the optimal incentive plan under centralization, and (ii) induces optimal effort and investment decisions.

REFERENCES

Anderson, M., R. Banker, and S. Ravindran. 2000. Executive compensation in the information technology industry. *Management Science* 46 (4): 530–547.

Antle, R., and G. Eppen. 1985. Capital rationing and organizational slack in capital budgeting. *Management Science* 31 (2): 163–174.

———, and J. Fellingham. 1990. Resource rationing and organizational slack in a two-period model. *Journal of Accounting Research* 28 (2): 1–24.

Baldenius, T. 2001. Delegated investment decisions and private benefits of control. Working paper, Columbia University.

Baron, D., and R. Myerson. 1982. Regulating a monopolist with unknown costs. *Econometrica* 50 (4): 911–933.

Battey, J. 2000. The keys to recruiting success. *InfoWorld* 22: 73–76.

Bernardo, A., H. Cai, and J. Luo. 2001. Capital budgeting and compensation with asymmetric information and moral hazard. *Journal of Financial Economics* 61 (3): 311–344.

Bhide, A. 2000. *The Origin and Evolution of New Businesses*. New York, NY: Oxford University Press.

Bushman, R., and R. Indjejikian. 1993. Accounting income, stock price, and managerial compensation. *Journal of Accounting and Economics* 16 (1–3): 3–23.

Christensen, P., G. Feltham, and M. Wu. 2002. Cost of capital in residual income for performance evaluation. *The Accounting Review* 77 (1): 1–23.

Core, J., and W. Guay. 2001. Stock option plans for non-executive employees. *Journal of Financial Economics* 61 (2): 207–225.

Datar, S., S. Kulp, and R. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39 (1): 75–92.

Dutta, S., and S. Reichelstein. 2002. Controlling investment decisions: Depreciation and capital charges. *Review of Accounting Studies* 7 (2–3): 253–281.

Ely, K. 1991. Interindustry differences in the relation between compensation and firm performance variables. *Journal of Accounting Research* 29 (1): 37–58.

Feltham, G., and D. Xie. 1994. Performance measure congruity and diversity in multitask principal-agent relations. *The Accounting Review* 69 (3): 429–454.

Harris, M., and A. Raviv. 1996. The capital budgeting process: Incentives and information. *Journal of Finance* 51 (4): 1139–1173.

Indjejikian, R., and D. Nanda. 1999. Dynamic incentives and responsibility accounting. *Journal of Accounting and Economics* 27 (2): 177–201.

Ittner, C., and D. Larcker. 1998. Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research* 6 (2): 205–238.

———, R. Lambert, and D. Larcker. 2001. The structure and performance consequences of equity grants to employees of new economy firms. Working paper, University of Pennsylvania.

Lambert, R. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32 (1–3): 3–87.

Lewis, T., and D. Sappington.1989. Countervailing incentives in agency problems. *Journal of Economic Theory* 49 (2): 294–313.

Portnot, S., and E. Moltzen. 2000. High-tech executives chase the dot-com dream. *Computer Reseller News* 895: 80.

Rajan, R., and L. Zingales. 2000. The governance of the new enterprise. In *Corporate Governance: Theoretical and Empirical Perspectives*, edited by X. Vives, 201–227. Cambridge, U.K.: Cambridge University Press.

Reichelstein, S. 1997. Investment decisions and managerial performance evaluation. *Review of Accounting Studies* 2 (4): 157–180.

Rogerson, W. 1997. Inter-temporal cost allocation and managerial investment incentives. *Journal of Political Economy* 105 (4): 770–795.

Stewart, B. 1991. *The Quest for Value*. New York, NY: Harper Collins Publishers.

Zingales, L. 2000. In search of new foundations. *The Journal of Finance* 55 (4): 1623–1653.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 95–117

Audit Committee Characteristics and Auditor Dismissals following "New" Going-Concern Reports

Joseph V. Carcello
University of Tennessee
Terry L. Neal
University of Kentucky

ABSTRACT: One important role of audit committees is to protect external auditors from dismissal following the issuance of an unfavorable report. We examine auditor dismissals following *new* going-concern reports that Big 6 firms issued between 1988 and 1999. Our findings suggest that audit committees with greater independence, greater governance expertise, and lower stockholdings are more effective in shielding auditors from dismissal after the issuance of new going-concern reports. In addition, we find that the relation between audit committee independence and auditor protection from dismissal has grown stronger over time. Finally, independent audit committee members experience a significant increase in turnover rate after auditor dismissals. These findings, coupled with those from Carcello and Neal (2000), suggest that when affiliated directors dominate the audit committee, management often can (1) pressure its auditor to issue an unmodified report despite going-concern issues, and (2) dismiss its auditor if the auditor refuses to issue an unmodified report.

Keywords: *audit committee characteristics; auditor dismissals; going-concern reports.*

Data Availability: *The data are available from public sources. A list of sample firms is available from the second author.*

We thank Mark Beasley, Dana Hermanson, Linda McDaniel, Mark Nelson (the associate editor), Bob Ramsay, two anonymous reviewers, and workshop participants at the University of Georgia, University of Oklahoma, and Louisiana State University for their helpful comments. We also thank Mike Barnhisel, Mike Borth, Becky Cassill, Marie Dothard, and especially Sandra Holt for research assistance.

Submitted April 2002
Accepted August 2002

I. INTRODUCTION

The sudden collapse of Enron amid questionable accounting practices has led Congress and regulators to call for more effective audit committee performance as one means of enhancing external auditor independence (Pitt 2001; Ruder 2002). This study contributes to the growing literature on corporate governance by investigating the relation between audit committee characteristics (independence, governance expertise, financial expertise, and stock ownership[1]) and auditor dismissals following the issuance of *new* going-concern reports. We define a going-concern report as new if the client received an unmodified (clean) report in the previous year.[2]

One of the primary functions of an audit committee is to safeguard the independence of the external auditor. Auditor independence is essential to audit quality because it minimizes "the possibility that any external factors will influence an auditor's judgments" (SEC 2000, 5). The independence of the auditor is particularly critical in financial accounting and reporting situations that are ambiguous, such as when a client is experiencing financial distress and the "temptation to 'see it the way your client does' is subtle, yet real" (Levitt 2000, 5). In their examination of an ambiguous reporting situation, Carcello and Neal (2000) find that audit firms are less likely to issue going-concern reports to financially distressed clients whose audit committees lack independence. Auditors may hesitate to issue a going-concern report if management implicitly or explicitly suggests that the client will dismiss the auditor if the auditor issues a going-concern report.

Prior research finds that clients receiving a going-concern report are more likely to switch auditors (e.g., Chow and Rice 1982; Mutchler 1984; Geiger et al. 1998). We test the hypothesis that management is less likely to terminate the auditor following the issuance of a going-concern report if the audit committee—which reviews all auditor-management disputes—embodies certain characteristics (PricewaterhouseCoopers 2000).

As shareholder representatives, the audit committee plays an important role in the auditor dismissal process. The NYSE-NASDAQ-sponsored Blue Ribbon Committee (BRC) (1999, 14) recognized the audit committee's "ultimate authority and responsibility to select, evaluate, and where appropriate, dismiss the outside auditor." Moreover, Securities and Exchange Commission Chairman Harvey Pitt has recently proposed additional safeguards to prevent management from firing the auditor without audit committee approval (Pitt 2002b).

Audit committee performance should be of high quality when members are independent (Public Oversight Board [POB] 1994; BRC 1999; PricewaterhouseCoopers 2000), when they have more governance expertise (Fama 1980; Fama and Jensen 1983), and more financial expertise (BRC 1999). Since critics have alleged that higher levels of stock ownership motivate corporate directors (and management) to artificially boost reported performance (Millstein 2002; Pitt 2002a), we expect that owning large stockholdings in the company will impair audit committee members' performance. In sum, we expect that audit committees whose members are more independent, have more governance expertise, more financial expertise, and own less of the company's stock, will be more likely to resist managerial attempts to dismiss an auditor following the issuance of a going-concern report.

Our results generally support our expectations. Auditors who issue a going-concern report are more likely to be dismissed if audit committees have a larger percentage of affiliated directors on the audit committee, or if audit committee members: (1) have less governance expertise, or (2) own a larger percentage of the company's stock. We do not find a significant relation between the percentage of audit committee members with financial expertise and auditor dismissals following going-concern reports.

---

[1] Throughout this paper, when we refer to the percentage of stock owned by audit committee members, we include their stock options.

[2] Subsequent references to going-concern reports denote new going-concern reports.

These results, coupled with the evidence reported by Carcello and Neal (2000), indicate potential problems in the interactions among auditors, audit committees, and management of financially distressed companies. Carcello and Neal (2000) suggest that auditors often believe that they are more likely to be dismissed following a going-concern report if there is a greater percentage of affiliated directors on the audit committee. The current study provides explicit evidence that this concern is valid. Therefore, we conclude that when affiliated directors dominate the audit committee, management often can (1) pressure its auditor to issue an unmodified report despite going-concern issues, and (2) dismiss its auditor if the auditor refuses to issue an unmodified report.

We organize the remainder of this paper as follows. Section II provides further background on the link between audit committee characteristics and auditor dismissals and develops our empirical predictions. We present the research design and sample selection procedure in Section III, and our results in Section IV. Section V contains supplemental analyses, and the last section discusses the study's implications and limitations.

II. BACKGROUND AND EMPIRICAL PREDICTIONS

Previous research suggests that opinion shopping is generally unsuccessful (e.g., Chow and Rice 1982; Smith 1986; Krishnan and Stephens 1995; Geiger et al. 1998). Nonetheless, clients receiving a going-concern report are more likely to switch auditors (e.g., Chow and Rice 1982; Mutchler 1984; Geiger et al. 1998) perhaps because management *believes* that once an incumbent auditor is dismissed, the company will find a more pliable auditor (Craswell 1988, 26). Alternatively, management might dismiss an auditor solely as punishment for issuing a going-concern report, or due to irreparable damage to its relationship with the auditor as a result of the conflict.

We posit that an audit committee whose members have more independence, governance expertise, financial expertise, and who own less of the company's stock will be more likely to block a managerial attempt to dismiss an auditor who issued a going-concern report. We now consider in more detail why we expect each of these four audit committee characteristics to affect auditor dismissals.

In our view, the independence of the audit committee is the primary mechanism for reducing the likelihood that the company will dismiss its auditor in retaliation for issuing a going-concern report. We classify audit committee members as either independent or affiliated directors. *Affiliated directors* (who have strong economic or personal ties to the company or its management) include current or former officers or employees of the company or of a related entity, relatives of management, professional advisors to the company (e.g., consultants, bank officers, legal counsel), officers of significant suppliers or customers of the company, and interlocking directors (Vicknair et al. 1993; Beasley 1996; Carcello and Neal 2000).

Given the strong economic or personal ties between affiliated directors and the company or its management, affiliated audit committee members have more incentive to side with management in disputes with the auditor (Baysinger and Butler 1985). Affiliated committee members also typically own more stock in the company than independent members, and stock ownership may increase the likelihood that a director sides with management. Indeed, we expect independent directors to be more objective and to possess greater expertise than affiliated directors, and the legal system appears to hold independent directors to a higher standard (Braiotta 1999, 104).

An audit committee with greater *governance expertise* should also reduce the likelihood that the company will dismiss its auditor in retaliation for issuing a going-concern report. Fama (1980) and Fama and Jensen (1983) suggest that directors make costly investments to develop reputations as effective monitors of corporate performance. Moreover, prior research concludes that directors of companies experiencing adverse events such as poor performance or financial distress subsequently serve less often as directors for other companies (Gilson 1990; Kaplan and Reishus 1990). We expect a director's reputation to suffer when the company fires its auditor after issuing a going-concern report. Directors who sit on more boards will have more to lose and therefore we expect them to be more likely to oppose the auditor's dismissal.

An audit committee with greater *financial expertise* should reduce the likelihood that the company will dismiss its auditor for issuing a going-concern report. The BRC (1999, 25) recommends that every audit committee have at least one member who possesses financial expertise, defined as "past employment experience in finance or accounting, requisite professional certification in accounting, or any other comparable experience or background which results in the individual's financial sophistication, including being or having been a CEO or other senior officer with financial oversight responsibilities." We expect that a financial expert will understand and support an auditor's decision to issue a going-concern report, and that an audit committee whose members have greater financial expertise will be more effective in preventing management from dismissing its auditor in this event.

An audit committee with a lower level of *stock ownership* should reduce the likelihood that the company dismisses its auditor after receiving a going-concern report. Audit committee members who own more company stock are likely to suffer losses if the going-concern report triggers a negative stock price response (Jones 1996; Melumad and Ziv 1997), so we expect them to be more willing to accede to auditor dismissals.[3]

III. RESEARCH DESIGN AND SAMPLE

Model

We use the following logistic regression model to test the relation between the likelihood that the client dismisses its auditor and audit committee characteristics:

$$\begin{aligned}\text{DISMISSED} = {} & b_0 + b_1\text{AFFILIATED} + b_2\text{AFFILIATED} \times \text{GC\_OPINION} \\ & + b_3\text{GOVEXPERT} + b_4\text{GOVEXPERT} \times \text{GC\_OPINION} \\ & + b_5\text{FINEXPERT} + b_6\text{FINEXPERT} \times \text{GC\_OPINION} \\ & + b_7\text{STOCKOWN} + b_8\text{STOCKOWN} \times \text{GC\_OPINION} + b_9\text{SIZE} \\ & + b_{10}\text{INDSHARE} + b_{11}\text{TENURE} + b_{12}\text{ZFC} + b_{13}\text{MGMTCHG} + \varepsilon.\end{aligned}$$

We define the dependent, test, and control variables as follows:

---

[3] Agency theory (e.g., Jensen and Meckling 1976) would suggest an opposite relation because a higher level of stock ownership should align the interests of management (audit committee members) with the interests of stockholders. Stock ownership is most likely to align the interests of audit committee members and stockholders if audit committee members are *long-term investors*. However, many audit committee members (and management) have a short-term perspective with respect to their ownership stake (Leonhardt 2002), and financial distress is likely to increase this myopia because the company's long-term existence is in doubt.

DISMISSED = identifies whether a client dismissed its auditor before the client's next annual report (1 = client dismissed auditor, 0 = client did not dismiss auditor);
AFFILIATED = the percentage of audit committee members classified as affiliated directors: current or former officers or employees of the company or of a related entity, relatives of management, professional advisors to the company, officers of significant suppliers or customers of the company, and interlocking directors;
GOVEXPERT = a proxy for audit committee members' governance expertise, the average number of directorship positions they hold in other public companies;
FINEXPERT = the percentage of audit committee members possessing financial expertise, per the BRC (1999) recommendations;[4]
STOCKOWN = the percentage of the client's common stock (and stock options) held by its audit committee members; and
GC_OPINION = a going-concern opinion indicator variable (1 = GC, 0 = clean).

Although we do not hypothesize significant relations between audit committee characteristics and auditor dismissals following a clean opinion, our analysis includes dismissals after both types of opinions. As explained later, we use a matched-pairs design where companies dismissing their auditor are matched on opinion, year, size, and industry, with companies not dismissing their auditor. This design allows us to determine the incremental relation between audit committee characteristics and dismissals following a going-concern (GC) opinion (relative to a clean opinion), which, in turn, enables us to address one potential alternate explanation for our results—audit committees with certain characteristics could be associated with auditor dismissals in all contexts, not just when management wants to dismiss the auditor because of a GC report.

The coefficients associated with the main effects of each of the four audit committee characteristics (b_1, b_3, b_5, b_7) measure the relation between the characteristic and auditor dismissals following clean opinions. However, we are primarily interested in the relation between audit committee characteristics and auditor dismissals following GC opinions. For each audit committee characteristic, we sum the coefficient on the main effect of that characteristic plus the coefficient on the term that interacts that characteristic with the going-concern opinion indicator variable.[5]

We expect that the higher the percentage of affiliated directors on the audit committee, the more likely the client is to dismiss its auditor following the receipt of a GC report ($b_1 + b_2 > 0$). We expect that the more directorship positions audit committee members hold, the less likely the client is to dismiss its auditor ($b_3 + b_4 < 0$). We also expect that the higher the percentage of audit committee members with financial expertise, the less likely the client is to dismiss its auditor ($b_5 + b_6 < 0$). Finally, we expect that the more of the client's common stock and options audit committee members hold, the more likely the client is to dismiss its auditor ($b_7 + b_8 > 0$).

---

[4] The BRC (1999) recommends that an audit committee contain at least one financial expert. As 95.2 percent of the audit committees in our study had at least one member with financial expertise, we measured financial expertise as the percentage of audit committee members possessing financial expertise as defined by the BRC (1999), rather than using a dichotomous measure.

[5] Because we use a matched-pairs design, including the base level of GC_OPINION as a main effect in the model would provide no incremental explanatory power (the dismissal rate is 50 percent for both the GC and clean samples).

In addition to the audit committee characteristics of interest, we control for the effects of other factors that likely affect a client's decision to dismiss its auditor: client size, auditor industry share, auditor tenure, financial distress, and managerial change.

SIZE. Large clients are less likely to dismiss their auditors (Francis and Wilson 1988; Haskins and Williams 1990; Krishnan 1994). Large clients wield more bargaining power as a result of the higher audit fees that they pay (McKeown et al. 1991, 11), and thus we expect fewer disagreements leading to auditor dismissals. Larger companies also have more incentive than smaller companies to retain their auditors, since financial analysts and the financial press scrutinize their auditor dismissals more closely. We measure size as the natural log of total assets (expressed in thousands of dollars). We expect that larger clients are less likely to dismiss their auditor.

INDSHARE. Prior studies provide evidence that a client is less likely to dismiss an auditor who specializes in the client's industry (Haskins and Williams 1990; Williams 1988). Williams suggests that clients commonly perceive that auditors with a greater market share in the industry are likely to be more effective, so clients would not want to change to a less effective auditor (i.e., an auditor with a lower market share). We measure an auditor's industry share as the percentage of the square root of total assets that the auditor audits for all companies in the client's industry (Hogan and Jeter 1999).[6] We expect that the larger the auditor's market share in the client's industry, the less likely the client is to dismiss its auditor.

TENURE. Levinthal and Fichman (1988) and Williams (1988) provide evidence that longer auditor tenure reduces the likelihood of auditor dismissal. As the tenure of the auditor-client relationship lengthens, the auditor develops client-specific knowledge that would be lost if the client changed auditors. We measure auditor tenure as the number of consecutive years that the client has retained the auditor.[7] We expect that the longer the auditor's tenure, the less likely the client is to dismiss its auditor.

ZFC. The financial condition of a company affects the likelihood that it will change auditors (Krishnan and Stephens 1995; Schwartz and Menon 1985). We compute the Zmijewski financial condition (ZFC) score using the PROBIT coefficients from Zmijewski's (1984, 69) 40 bankrupt/800 nonbankrupt estimation sample.[8] Higher ZFC scores indicate greater financial distress, so we expect that the higher the ZFC score, the more likely the client is to dismiss its auditor.

MGMTCHG. A change in top management is often associated with a change in auditors (Chow and Rice 1982; Williams 1988). A new CEO could change auditors to obtain a fresh perspective on the company's financial results, or because he or she had positive experiences with another audit firm. A new chief financial officer (CFO) could seek an auditor change for similar reasons, and CFOs with prior experience in public accounting often prefer to work with their former employers (Iyer 1998). We use a dummy variable (1 = CEO or CFO change, 0 = no change) to measure whether top management changed in the year the opinion was issued or in the following year, provided the change in top management preceded the date of the auditor dismissal.[9] We expect a change in the company's CEO or CFO to increase the likelihood that the client dismisses its auditor.

---

[6] We based industry share measures on three-digit SIC codes for all industries with at least ten companies. For the 21 industries with fewer than ten companies, we determined the industry share measures at the two-digit level.

[7] We truncate auditor tenure at ten years to reduce the effect of extreme values for clients that have retained their auditors for many years.

[8] A number of prior studies have used Zmijewski's (1984) ZFC score to control for financial distress level (e.g., Bamber et al. 1993; Carcello et al. 1995).

Sample

We identified all publicly traded companies (except financial institutions and service companies [i.e., SIC codes 6000–9999]) included on the Compact D/SEC database that (1) received a new GC report from a Big 6 firm during the period 1988 to 1999, and (2) dismissed their auditor prior to the issuance of the client's next annual report. We limit our sample to companies with SIC codes below 6000 because Zmijewski (1984) excluded financial and service companies when he developed his financial condition index (ZFC), which is our control variable for client financial condition. Although auditor changes include dismissals and resignations, we limit our study to auditor dismissals because we are interested in whether audit committee characteristics enhance auditor independence by making it less likely that the client dismisses its auditor, not whether audit committee characteristics make it less likely that the auditor dismisses its client.

One hundred seventy-four companies met our sample selection criteria. We then exclude companies that filed for bankruptcy during the fiscal year under audit or the next year ($n = 30$), because the bankruptcy itself sometimes precipitates the auditor dismissal (Williams 1988; DeFond 1992; Krishnan 1994). Since we focus on the relation between audit committee characteristics and auditor dismissals, we exclude companies that did not maintain an audit committee ($n = 45$). We also exclude companies for which we were unable to obtain the relevant proxy statement or Form 10-K ($n = 33$), from which we gathered data on audit committee characteristics.

We then matched each GC client that dismissed its auditor with a GC client that did not dismiss its auditor.[10] We matched on year, industry, and to the extent possible, size.[11] We were unable to locate a suitable matching client for four GC dismissal companies because the number of GC companies that dismissed their auditors in that industry and year exceeded the number of potentially matching GC companies that did not dismiss their auditors. Table 1 shows that this resulted in a final sample of 62 clients receiving GC reports that dismissed their auditors, matched with 62 clients receiving GC reports that retained their auditors.

To provide a benchmark for assessing the relation between audit committee characteristics and auditor dismissals, we gathered a sample of companies that received clean opinions between 1988 and 1999 and dismissed their auditors during the following year. Approximately 1,000 public companies (with SIC codes below 6000) received clean reports from Big 6 firms between 1988 and 1999 and then changed their auditors during the following year. Due to the large sample size, we selected a random sample of 200 companies. From this sample, 24 companies did not maintain audit committees and 13 other companies did not have financial statement or other required data.

---

[9] For companies that did not change auditors, we determined whether top management changed during the year corresponding with the financial statements, or in the first nine months of the next year.

[10] The nondismissal category includes two clients that changed auditors in the second year after the GC opinion. Excluding these clients from our analyses does not affect our inferences.

[11] We matched 13 clients at the four-digit SIC code level, 16 at the three-digit level, 16 at the two-digit level, and 17 at the one-digit level. We then selected the GC client retaining its auditor that was closest in total assets to the GC client that dismissed its auditor. Given the limited number of GC observations, we chose not to use a stricter match on size (e.g., match on assets within ±20 percent).

TABLE 1
Sample Selection Criteria

	Going-Concern Opinion Sample[a]	*Clean Opinion Sample*[b]	*Combined Sample*
Initial sample	174	200	374
Companies that filed for bankruptcy	(30)		(30)
Companies with no audit committee	(45)	(24)	(69)
Proxy statement and/or Form 10-K not available	(33)	(13)	(46)
Disagreement over accounting or auditing issues		(7)	(7)
Auditor resigned		(14)	(14)
No companies meeting matching criteria existed	(4)	(17)	(21)
Dismissal sample	62	125	187
Nondismissal sample[c]	62	125	187
Total sample	124	250	374

[a] The going-concern opinion sample frame is publicly held companies that dismissed their auditors after receiving *new* going-concern reports from Big 6 audit firms during the period 1988–1999 (financial institutions and service companies excluded).

[b] The clean opinion sample frame is publicly held companies that dismissed their auditors after receiving clean reports from Big 6 audit firms during the period 1988–1999 (financial institutions and service companies excluded). Due to the large sample size, we selected an initial random sample of 200 companies.

[c] We matched every company in the sample that dismissed its auditor with a company that received the same type of opinion but did not dismiss its auditor. We matched companies on year, industry, and to the extent possible, size.

We are particularly interested in controlling for the relation between audit committee characteristics and auditor dismissals that do not result from a conflict between the auditor and management (i.e., benign auditor changes). We are interested in benign changes because we would expect the same relation between audit committee characteristics and auditor dismissals following an auditor-client disagreement over accounting and auditing issues as between audit committee characteristics and dismissals following a GC opinion. Including non-GC contentious auditor dismissals in the clean opinion sample would blur the distinction between the GC and clean opinion samples. (Unfortunately, given the limited number of auditor dismissals following an accounting/auditing disagreement [n = 7], we cannot test the relation between audit committee characteristics and non-GC contentious auditor dismissals.)

We examined the 8-Ks filed after each auditor change to determine the reasons the client gave for changing auditors. We excluded observations where the client and auditor disagreed over accounting or auditing issues (n = 7) and where the auditor resigned (n = 14). We then matched each company that received a clean opinion and dismissed its auditor with a company that received a clean opinion and did not change its auditor. We matched on year, industry, and size. We could not match on size for 17 of the dismissal companies in the clean opinion sample.[12] As shown in Table 1, our final sample includes

---

[12] Of our 125 matched pairs, we matched 61 at the four-digit SIC code level, 20 at the three-digit level, and 44 at the two-digit level. We matched on size using total assets, if possible, within ±20 percent. For ten matched-pairs, the size match was ±30 percent.

125 clients that received a clean opinion and dismissed their auditor, matched with 125 companies that received a clean opinion and did not change their auditor.

Therefore, our final sample includes 250 companies receiving clean opinions and 124 companies receiving GC opinions. By sample construction, 50 percent of the companies receiving each type of opinion dismissed their auditor.

IV. RESULTS

Univariate Analysis

Table 2 presents descriptive statistics for the audit committee characteristics and control variables. Since we are primarily interested in the relation between audit committee characteristics and dismissals following a GC report, we present separate descriptive statistics for the audit committee characteristics for clients receiving GC and clean reports. As we expect, three of the four characteristics differ significantly between clients dismissing their auditor following a GC opinion and those clients retaining their auditor after receiving a GC opinion. For GC clients who dismissed their auditors, 49 percent of the audit committee are affiliated directors, whereas for GC clients that did not dismiss their auditor, only 24 percent of the audit committee are affiliated directors ($p < 0.01$). GC clients that dismissed their auditors had audit committee members who held an average of only 0.84 directorships in other companies, while GC clients who retained their auditors had audit committee members holding an average of 1.83 directorships ($p < 0.01$). Audit committee members' average percentage stock ownership for GC clients dismissing their auditors is 9 percent compared to 3 percent for GC clients retaining their auditors ($p < 0.01$). There is no significant difference in the percentage of audit committee members with financial expertise between those clients dismissing their auditors vs. those clients retaining their auditors. In addition, there are no significant differences for any of the audit committee characteristics for clients dismissing their auditors following a clean opinion vs. clients retaining their auditors after receiving a clean opinion.

With respect to the control variables, clients are more likely to dismiss auditors when there has been a recent change in the client's top management ($p < 0.05$). There is no significant relation between auditor dismissals and auditor industry-market share, auditor tenure, and client financial distress level. In addition, there is no significant relation between client size[13] and auditor dismissals in our sample, although this result is not surprising given that we matched on client size.

Table 3 presents the correlations among the independent variables. The correlations are all below 0.35.[14] A higher percentage of affiliated directors on the audit committee is associated with more stock ownership ($\text{rho} = 0.29$, $p < 0.01$). In results not separately reported, we find that affiliated audit committee members on average own 3.3 percent of their companies' shares, whereas independent audit committee members own only 1.0 percent. As discussed previously, affiliated directors' higher stock ownership can increase their incentive to support managerial efforts to replace an auditor issuing a GC report.

In addition, audit committees with a higher percentage of affiliated directors are on average smaller ($\text{rho} = -0.19$, $p < 0.01$), have members with less governance expertise ($\text{rho} = -0.21$, $p < 0.01$), and experience more financial distress ($\text{rho} = 0.12$, $p < 0.05$).

---

[13] Although we control for client size in our logistic regression model using the natural log of total assets, we present descriptive statistics on client size for total assets (untransformed).

[14] The generally modest correlations suggest that multicollinearity is unlikely to be a problem, but we nonetheless computed variance inflation factors. Only two of the variance inflation factors are above 3.0 and the largest is 3.7; Gujarati (1995, 339) suggests that multicollinearity is unlikely to be problematic if the variance inflation factors are below 10.0.

TABLE 2
Descriptive Statistics for Audit Committee Characteristics and Control Variables
Mean (Median) [Standard Deviation]

	Combined Sample			*Going-Concern Sample*			*Clean Sample*		
Variable[a]	*Dismissal (n = 187)*[b]	*No Dismissal (n = 187)*[c]	*(Dismissal − No Dismissal)*[d]	*Dismissal (n = 62)*[b]	*No Dismissal (n = 62)*[c]	*(Dismissal − No Dismissal)*[d]	*Dismissal (n = 125)*[b]	*No Dismissal (n = 125)*[c]	*(Dismissal − No Dismissal)*[d]
Audit Committee Characteristics									
AFFILIATED	0.31	0.24	0.07*	0.49	0.24	0.25***	0.22	0.25	−0.03
	(0.33)	(0.00)	(0.33)**	(0.50)	(0.07)	(0.43)***	(0.00)	(0.00)	(0.00)
	[0.32]	[0.30]		[0.36]	[0.28]		[0.26]	[0.31]	
GOVEXPERT	1.02	1.49	−0.47***	0.84	1.83	−0.99***	1.11	1.32	−0.21
	(0.67)	(1.00)	(−0.33)***	(0.50)	(1.37)	(−0.87)***	(0.75)	(1.00)	(−0.25)
	[1.11]	[1.41]		[1.01]	[1.63]		[1.15]	[1.26]	
FINEXPERT	0.79	0.75	0.04	0.80	0.79	0.01	0.78	0.73	0.05
	(1.00)	(0.75)	(0.25)	(1.00)	(1.00)	(0.00)	(1.00)	(0.75)	(0.25)
	[0.27]	[0.28]		[0.30]	[0.28]		[0.25]	[0.28]	
STOCKOWN	0.05	0.03	0.02	0.09	0.03	0.06***	0.03	0.03	0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
	[0.11]	[0.07]		[0.15]	[0.07]		[0.07]	[0.08]	
Control Variables									
SIZE	352.74	334.26	18.48						
	(51.68)	(55.22)	(−3.54)						
	[1,065.39]	[995.75]							
INDSHARE	0.18	0.19	−0.01						
	(0.15)	(0.17)	(−0.02)						
	[0.11]	[0.11]							
TENURE	6.54	6.97	−0.43						
	(6.00)	(7.00)	(−1.00)						
	[3.07]	[3.04]							

(continued on next page)

Vicknair, D., K. Hickman, and K. C. Carnes. 1993. A note on audit committee independence: Ev from the NYSE on "grey" area directors. *Accounting Horizons* 7 (March): 53–57.

Williams, D. D. 1988. The potential determinants of auditor change. *Journal of Business Fina Accounting* 15 (Summer): 243–261.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress pre models. *Journal of Accounting Research* 22 (Supplement): 59–82.

THE ACCOUNTING REVIEW
Vol. 78 No. 1
January 2003
pp. 119–142

The Effects of Debt Contracting on Voluntary Accounting Method Changes

Anne Beatty
Pennsylvania State University
Joseph Weber
Massachusetts Institute of Technology

ABSTRACT: This study examines whether the provisions of a firm's bank debt contracts affect its accounting choices. Starting with a sample of firms who have bank debt and who also voluntarily changed accounting methods, we investigate whether the likelihood that the change in accounting method increased (rather than decreased) the borrower's income depends on (1) whether the change in accounting method affects the bank debt contract calculations, (2) the expected costs of violating the bank debt covenants, (3) whether performance pricing provisions affect the interest rate on the loan, and (4) whether the bank debt contract contains accounting-based dividend restrictions. After controlling for other motives for changing accounting methods, we find that borrowers whose bank debt contracts allow accounting method changes to affect contact calculations are more likely to make income-increasing rather than income-decreasing changes. This increase in likelihood of an income-increasing change is attenuated when expected costs of technical violation are lower because there is a single lender, and occurs for borrowers whose debt contacts have performance pricing and dividend restrictions. These results suggest that incentives to lower interest rates through performance pricing or to retain dividend payment flexibility influence borrowers' accounting method choices, thereby addressing the fundamental questions posed by Fields et al. (2001) of whether, under what circumstances, and how accounting choice matters.

Keywords: *accounting choice; debt contracting; debt covenants; performance pricing; dividend restrictions.*

We thank two anonymous reviewers, Wayne Guay, Peter Joos, S. P. Kothari, Greg Miller, and seminar participant at MIT for helpful comments. We also thank Jennifer Altamuro, Shelley Herbein, and Hyunna Song for research assistance and Dick Dietrich for his insights about data availability at the SEC. Professor Beatty gratefully acknowledges financial support from PricewaterhouseCoopers and the Sloan Foundation.

Submitted February 2002
Accepted August 2002

Data Availability: *All data are available from public sources.*

I. INTRODUCTION

Watts and Zimmerman (1986, 216–217) argue that debt contracts that make covenant thresholds a function of financial ratios give borrowers an incentive to change accounting methods to avoid costly covenant violations. However, reviewing the empirical research on accounting choice, Fields et al. (2001, 275) state that "the evidence on whether accounting choices are motivated by debt covenant concerns is inconclusive." There are at least three reasons that previous research may underestimate the effects of debt contracting on managers' accounting choices.

First, previous research that does not use the details of firms' actual debt contracts assumes that contract calculations are based on current accounting methods. However, Mohrman (1996) and Beatty, Ramesh, and Weber (2002) document that private debt contract calculations often prohibit firms from changing accounting methods. For these contracts, borrowers cannot use voluntary accounting method changes to avoid covenant violations. Thus, debt contracts give borrowers an incentive to change accounting methods only if the voluntary accounting method changes affect contract calculations.

Second, existing research that focuses exclusively on debt covenants ignores other accounting-based features of debt contracts. Performance pricing is a relatively new feature in bank debt contracts that explicitly makes the interest rate charged on a bank loan a function of the borrower's current creditworthiness. Asquith et al. (2002) document that performance-pricing features typically measure the borrower's creditworthiness using financial ratios such as debt to earnings before interest, taxes, depreciation, and amortization (EBITDA), leverage, or interest coverage. That is, the interest rate charged in the contract does not remain fixed over the length of the loan, but varies inversely with changes in measures of financial performance. Compared to covenants under which accounting information affects loan rates only when the borrower violates a single threshold, performance pricing creates a more continuous and direct link between accounting information and interest rates. Thus, performance pricing likely gives managers additional incentives to make income-increasing accounting method changes. Previous research does not consider whether this feature of debt contracts influences accounting choice.

Third, previous research has focused on borrowers who were either close to violating or had already violated covenants. These studies may underestimate the effect of debt contracting on accounting choice for cases in which borrowers have effectively used accounting changes to provide slack in financial covenants, and thus never come close to covenant violations.

To provide more conclusive evidence that debt contracting concerns affect borrowers' accounting choices, we address these three limitations of prior research. First, our research design incorporates the fact that debt contracts often prohibit borrowers from using voluntary accounting method changes to affect contract calculations. This cross-sectional variation in bank debt provisions allows us to examine whether managers who have chosen to change their accounting methods are more likely to make income-increasing changes when the debt contract allows these changes to affect contract calculations. Second, we consider whether another debt contracting feature—accounting-based performance pricing—increases borrowers' tendencies to make their voluntary accounting method changes income-increasing rather than income-decreasing. Third, rather than restricting our sample to borrowers approaching covenant violations, we study all borrowers who make voluntary accounting changes, and control for their other incentives to change accounting methods. If borrowers effectively use changes in accounting methods to create slack in financial covenants so that they avoid coming close to violating their covenants, then our tests examining all accounting method changes are more powerful than tests employed in previous research that focused solely on borrowers who are close to or had already violated their covenants.

To provide evidence on how debt contracts affect the sign of voluntary changes that borrowing firms make in their accounting methods, we identify a sample of borrowers with material bank debt who have chosen to make accounting changes. We determine whether their contracts have accounting-based restrictive covenants, dividend restrictions, or performance-pricing requirements. Beatty, Ramesh, and Weber (2002) find that borrowers willingly pay higher interest rates to obtain bank debt contracts that allow voluntary accounting changes to affect the calculation of financial terms in those debt contracts. We therefore expect borrowers to make use of this costly flexibility. Consistent with our hypothesis, we find that borrowers who change their accounting methods are more likely to make income-increasing changes if their debt contracts allow the changes to affect contract calculations. This increase in likelihood of an income-increasing accounting change is attenuated when the cost of a covenant violation is lower because all the firm's bank debt is from a single lender, and occurs only for borrowers whose contracts have performance-pricing provisions and dividend restrictions.

Our debt contracting results hold even after we control for other motives that borrowers have for changing accounting methods. Specifically, our results hold even after we control for the fact that incentives arising from executive compensation contracts and incentives to meet earnings benchmarks increase borrowers' propensity to make income-increasing (rather than income-decreasing) changes. We also control for the fact that a borrower with a new Chief Executive Officer (CEO) who reports a large loss before the effect of the accounting method change is more likely to report an income-decreasing accounting change (consistent with new CEOs of poorly performing firms taking "big baths"). Our debt contracting results also hold after controlling for tax incentives.

Our evidence that debt contracting concerns affect borrowers' accounting choices is more conclusive than previous research because we show that borrowers take advantage of the flexibility to make income-increasing accounting method changes that affect contract calculations. In addition, our evidence suggesting that incentives to lower interest rates through performance pricing and incentives to retain dividend payment flexibility appear to affect borrowers' accounting method choices helps address Fields et al.'s (2001) fundamental questions of whether, under what circumstances, and how accounting choice matters.

Section II develops our hypotheses and Section III explains our research design. Section IV describes our sample selection and provides descriptive statistics. We discuss our empirical results in Section V, provide sensitivity analysis in Section VI, and present our conclusions in Section VII.

II. HYPOTHESIS DEVELOPMENT

Fields et al. (2001) discuss three motives for accounting choice: (1) contracting (including debt and management compensation contracts), (2) asset pricing, and (3) influencing external parties (e.g., the Internal Revenue Service [IRS]). This paper focuses on the relation between accounting method changes and debt contracting while controlling for the other incentives for accounting choice, as explained in Section III. Specifically, we examine the extent to which detailed provisions of debt contracts explain the sign of borrowers' voluntary changes in accounting methods. We focus on accounting method changes because Beatty, Ramesh, and Weber (2002) document that borrowers willingly pay higher interest rates to retain the flexibility to make voluntary accounting method changes that affect bank debt contract calculations. We therefore expect borrowers to make use of this costly flexibility by making voluntary accounting method changes to affect contract calculations.

Does Allowing Accounting Changes to Affect Debt Contract Calculations Influence Accounting Choice?

Previous research examining whether debt covenants affect borrowers' decisions to change accounting methods generally assumes that the calculations stipulated in the covenants will be based on current accounting methods. For example, Healy and Palepu (1990) examine whether borrowers close to violating their dividend covenant restrictions, which are typically defined as a percentage of net income or retained earnings, change their accounting methods to increase those limits. Similarly, Sweeney (1994) investigates whether borrowers use accounting changes to avoid violating financial covenants designed to monitor borrowers' performance. However, Mohrman (1996) and Beatty, Ramesh, and Weber (2002) document that debt contract calculations are often based on the accounting rules that the borrower used when the contract originated. For these contracts, the borrower cannot use changes in accounting methods to avoid covenant violations. The cross-sectional variation in the effects of accounting method changes on contract calculations provides a natural setting that allows us to investigate the effects of debt contracting on borrowers' decisions to make accounting method changes.

Our first hypothesis is that borrowers who change accounting methods will be more likely to make income-increasing (rather than income-decreasing) accounting method changes when such changes affect covenant calculations.

Do the Expected Costs of Covenant Violations Influence Accounting Choice?

Watts and Zimmerman (1986, 215–216) argue that costly covenant violations give borrowers an incentive to make income-increasing accounting method changes to create slack in covenants; however, evidence on the actual magnitude of the cost of covenant violations is mixed. On the one hand, Beneish and Press (1993) provide evidence that covenant violations that lenders have not waived as of the financial-reporting date cost the borrower an average 80 basis point increase in the interest rate the lender charges on the loan after renegotiation. Furthermore, Healy and Palepu (1990) find that firms close to contractually imposed dividend limitations generally cut the dividends they pay rather than violate the dividend-restricting covenants or change their accounting methods to circumvent the covenants. These findings suggest that borrowers perceive violating dividend restriction covenants to be more costly than cutting dividends. On the other hand, Dichev and Skinner (2002) report that lenders may commonly waive technical violations prior to the financial reporting date, and they argue that waived covenant violations are unlikely to be very costly to the borrower. These results suggest that the cost of technical covenant violations will be lower the more likely it is for the lender to waive violations.

Our second hypothesis is that borrowers whose bank debt contracts allow voluntary accounting changes to affect contract calculations will be less likely to make income-increasing (rather than income-decreasing) accounting method changes when they expect lower technical default costs.

Does Performance Pricing Influence Accounting Choice?

We expect performance pricing, which is a relatively new debt-contracting feature, to give borrowers an additional incentive to make income-increasing accounting changes beyond the incentive provided by covenants. Accounting-based performance pricing explicitly makes the interest charged on a bank loan a function of the borrower's financial ratios, such as debt to EBITDA, leverage, or interest coverage. That is, the interest rate in the contract is not fixed over the length of the loan but varies inversely with changes in accounting measures of financial performance, which are typically assessed quarterly using the previous four quarters' results. Performance pricing may increase the importance of debt contracts in accounting choice by providing a more continuous and direct link between accounting information and interest rates than covenants provide.

Asquith et al. (2002) provide descriptive evidence on this new feature found primarily in bank debt contracts. They document that at the inception of the contract the interest rate charged on the loan is rarely above the top level or below the bottom level of the performance-pricing grid, that the average grid has five levels, and that the average increase in interest rates charged when the borrower's performance deteriorates one performance-pricing level is 15 basis points. Beatty, Dichev, and Weber (2002) document that a contract with performance pricing typically has a covenant based on the performance pricing ratio that is set at one level of performance below the worst level included in the grid. Because performance-pricing grids have multiple levels, deteriorations in a borrower's performance may result in a new performance-pricing level, and thus a higher interest rate, even if the borrower's performance never deteriorates enough to violate the covenant. In addition, improvements in a borrower's performance may result in a lower interest rate if the borrower reaches a better performance-pricing level. Thus, we expect performance pricing to provide an additional incentive for borrowers to make income-increasing accounting method changes.

Our third hypothesis is that borrowers with debt contracts allowing voluntary accounting method changes to affect contract calculations will be more likely to make income-increasing (rather than income-decreasing) accounting method changes if their contracts include accounting-based performance pricing in addition to traditional covenants.

Do Dividend Restrictions Influence Accounting Choice?

Borrowers' incentives to make income-increasing accounting method changes may also depend on whether their debt contracts contain dividend restrictions based on accounting performance. Debt contracts often have covenants that either explicitly limit dividend payments to be less than a given percentage of net income, or implicitly limit dividend payments by including covenants affected by dividends (e.g., net worth, tangible net worth, or leverage ratio covenants). The incentives to make income-increasing rather than income-decreasing accounting method changes provided by covenants that explicitly or implicitly restrict dividends may differ from those provided by other covenants (e.g., working capital or coverage ratio covenants). Healy and Palepu (1990) suggest that borrowers' incentive to make an income-increasing accounting change may be lower for covenants affected by dividend payments, because borrowers have the option to reduce dividends rather than violating the covenant. On the other hand, Sweeney (1994) documents that covenants affected by dividend payments are more likely than other covenants to be binding, suggesting that they may give borrowers a greater incentive to make an income-increasing accounting change. Thus, we are unable to predict the direction in which dividend restrictions will affect borrowers' accounting choices.

Our fourth hypothesis is therefore a nondirectional one: the likelihood that borrowers with debt contracts allowing voluntary accounting method changes to affect contract calculations make income-increasing (rather than income-decreasing) accounting method changes will depend on whether their debt contracts contain accounting-based dividend restrictions.

III. RESEARCH DESIGN

All of our tests use a sample of borrowers who have bank debt and who have also made voluntary changes in their accounting methods. Focusing on firms that make accounting changes has the advantage of requiring that the firms have enough flexibility to make an accounting change. Firms that want to make income-increasing accounting changes but are unable to do so should make no such changes and therefore not be included in our sample. This increases the power of our tests relative to prior research that examines the debt-contracting hypothesis using samples of borrowers who are close to violating covenants, which suffers from the problem that many of these borrowers may not have the flexibility to make accounting changes.

We focus on the effect of *bank debt* (rather than public debt) provisions on the borrower's voluntary accounting method changes for three reasons. First, Beatty, Ramesh, and Weber (2002) find that, in contrast to public debt, there is variation in whether bank debt contracts allow voluntary accounting changes to affect contract calculations. Second, El-Gazaar and Pastena (1991) find that bank debt contracts virtually always contain covenants, whereas public debt often does not, and that the covenants in bank debt are typically tighter than those in public debt.[1] Third, Asquith et al. (2002) find that, unlike public debt, many bank debt contracts have performance-pricing provisions, an innovation that likely provides additional incentives for firms to make income-increasing accounting changes. The greater cross-sectional variation in whether accounting changes affect contract calculations, combined with the potentially greater importance of covenants and performance pricing in bank debt compared to public debt, should increase the power of our tests.

Testing Hypotheses 1 and 2

We test our first hypothesis by examining whether a borrower's decision to make an income-increasing (rather than an income-decreasing) accounting method change depends on whether that change affects debt contract calculations. We test our second hypothesis by examining whether this relation depends on the expected costs of technical covenant violations. To control for other incentives for making income-increasing accounting method changes, such as those provided by management compensation contracts, asset pricing, and external parties like the IRS, as described by Fields et al. (2001), we estimate the following logistic regression:

$$\begin{aligned}\text{INCREASE} = \beta_0 &+ \beta_1\text{ACC\_Change} + \beta_2\text{One\_Lender} + \beta_3\text{COMP} + \beta_4\text{SMLOSS}\\ &+ \beta_5\text{SMLOSS\_NL} + \beta_6\text{LGLOSS} + \beta_7\text{LGLOSS\_}\Delta\text{CEO}\\ &+ \beta_8\text{NOL} + \beta_9\text{SIZE} + \varepsilon, \qquad (1)\end{aligned}$$

where:

Dependent Variable:

INCREASE = 1 if the accounting method change increases income and equity, 0 otherwise.

---

[1] Our examination of public debt and private placement agreements that are available online revealed that few of these contracts contained covenants or accounting-based performance pricing.

Debt-Contracting Independent Variables:
ACC_Change = 1 if contract allows accounting changes to affect calculations, 0 otherwise; and
One_lender = 1 if accounting method changes affect debt contract calculations *and* a single lender makes the loan, 0 otherwise.

Control Variables:
COMP = 1 if the firm indicates that the managers' bonuses depend on financial performance *and* the CEO received a bonus in the year prior to the accounting change, 0 otherwise;
SMLOSS = 1 if the firm's earnings before the effect of the accounting change is between 0 and −1 percent of assets, 0 otherwise;
SMLOSS_NL = 1 if the firm's earnings before the effect of the accounting change is between 0 and −1 percent of assets *and* the firm is not listed on a major stock exchange, 0 otherwise;
LGLOSS = 1 if the firm's earnings before the effect of the accounting change is less than −1 percent of assets, 0 otherwise;
LGLOSS_ΔCEO = 1 if the firm's earnings, before the effect of the accounting change, is less than −1 percent of assets *and* the firm's CEO changed in the year prior to the accounting change, 0 otherwise;
NOL = 1 for borrowers who have a net operating loss carryforward *and* make an accounting change that has a tax effect, 0 otherwise; and
SIZE = the log of the firm's assets before the effect of the accounting change.

Our dependent variable is a simple dichotomous variable that equals 1 if the accounting change increases the current period net income, and 0 if the accounting method change decreases net income.

Debt-Contracting Variables

To test our first hypothesis, our logit includes a dichotomous independent variable (ACC_Change) that equals 1 if the borrower has *any* bank debt contracts that allow accounting changes to affect contract calculations, 0 otherwise. To classify the effect of accounting changes on a borrower's contract calculations, we obtain copies of all of the bank debt contracts the borrower has entered at the time of the accounting change. We then use an approach similar to Mohrman (1996) and Beatty, Ramesh and Weber (2002) to identify whether accounting changes affect the contract calculations. Specifically, we classify a contract as *not* allowing voluntary accounting changes to affect calculations if the contract states either (1) that the borrower shall not make any significant change in accounting treatment or reporting practices, except as required by generally accepted accounting principles (GAAP) or (2) that accounting terms will be based on financial information prepared in accordance with GAAP as in effect and applied on the date of the agreement. If any of a borrower's bank debt contracts allow accounting changes to affect calculations, we then set ACC_Change equal to 1. Only when all of the firm's bank debt contracts prohibit accounting changes from affecting calculations do we set ACC_Change equal to zero.

To test our second hypothesis, we include a dichotomous independent variable (One_lender) equal to 1 if accounting method changes affect debt contract calculations *and* a single lender makes the loan, 0 otherwise. Asquith et al. (2002) argue that renegotiation costs are lower when there is a single lender, because contracts typically require a majority, supermajority, or unanimous consent of all lenders to waive a covenant. Therefore, the expected costs of technical default are lower when there is only one lender.[2]

Control Variables

To isolate the relation between accounting changes and debt contracting costs, we control for firms' other incentives to make accounting changes. In their summary of the empirical and theoretical accounting choice literature, Fields et al. (2001) suggest that management compensation contracts, asset pricing, and external parties such as the IRS provide additional incentives affecting firms' accounting choices beyond those provided by debt contracting.

To control for incentives arising from accounting-based compensation contracts, we include a dichotomous variable (COMP) equal to 1 if a firm indicates that its managers' bonuses depend on accounting-based measures *and* its CEO received a bonus in the year prior to the accounting change. We expect firms that pay accounting-performance-based bonuses to be more likely than other firms to make income-increasing (rather than income-decreasing) accounting method changes.

To capture borrowers' asset-pricing incentives to manage earnings, we control for borrowers' potential attempts to increase share prices by avoiding reporting a loss. Specifically, we include two variables to control for borrowers' use of accounting changes to achieve the positive earnings threshold Burgstahler and Dichev (1997) document. We set the first variable (SMLOSS) equal to 1 if the firm would have reported a loss of 1 percent of assets or less prior to the effect of the accounting change. These firms have an incentive to make an income-increasing accounting change to avoid reporting a loss. We construct the second variable (SMLOSS_NL) by interacting SMLOSS with a dichotomous variable equal to 1 for firms not listed on a major stock exchange, 0 otherwise. Based on the results reported in Beatty, Ke, and Petroni (2002), we expect firms not listed on a major exchange to have less incentive to avoid missing earnings benchmarks.

For firms expecting to report a large loss in the absence of an accounting change, management compensation and asset-pricing incentives may prompt managers to increase the magnitude of the loss to create reserves for the future. We include two variables to control for this "big bath" incentive. The first variable (LGLOSS) equals 1 if the firm would have reported a loss greater than 1 percent of assets prior to the effect of the accounting change. We construct the second variable (LGLOSS_ΔCEO) by interacting LGLOSS with a dichotomous variable equal to 1 if the firm changed CEOs in the year prior to the accounting change, 0 otherwise. We expect the big bath incentive to be higher for new CEOs who can use poor past performance to explain the magnitude of the loss.

To control for borrowers' tax incentives to change accounting methods, we include a dichotomous variable (NOL) equal to 1 for borrowers who have net operating loss carryforwards, and who also make accounting changes that have tax effects; 0 otherwise. We expect firms with NOL carryforwards to be more likely than others to make income-increasing accounting method changes that affect tax accounting.

Finally, we include the log of a firm's assets to control for firm size (SIZE). We make no prediction about the effect of firm size on the likelihood the manager makes an income-increasing rather than income-decreasing accounting method change.

---

[2] This hypothesis assumes that the covenants in contracts with single lenders are similar to the covenants in contracts with multiple lenders. We investigate the validity of this assumption, and find that the average number of covenants is virtually the same.

Testing Hypotheses 3 and 4

To test our third and fourth hypotheses, we re-estimate our model by substituting the specific type of debt contract feature affected by the accounting method changes for our ACC_Change variable. We partition firms whose contracts allow accounting changes to affect contract calculations into four mutually exclusive categories: those with neither performance pricing nor dividend restriction provisions; those with performance-pricing but no dividend restrictions; those with dividend restrictions but no performance-pricing; and those with both performance-pricing and dividend restrictions. Specifically, we estimate the following logistic regression:

$$\begin{aligned}\text{INCREASE} = {} & \beta_0 + \beta_{1a}\text{NoPPorDiv\_Res} + \beta_{1b}\text{PP} + \beta_{1c}\text{Div\_Res} \\ & + \beta_{1d}\text{PP\&Div\_Res} + \beta_2\text{One\_Lender} + \beta_3\text{COMP} + \beta_4\text{SMLOSS} \\ & + \beta_5\text{SMLOSS\_NL} + \beta_6\text{LGLOSS} + \beta_7\text{LGLOSS\_}\Delta\text{CEO} \\ & + \beta_8\text{NOL} + \beta_9\text{SIZE} + \varepsilon, \qquad (2)\end{aligned}$$

where:

NoPPorDiv_Res = 1 if the contract allows accounting changes to affect contract calculations but contains neither an accounting-based performance-pricing provision nor an accounting-based dividend restriction, 0 otherwise;

PP = 1 if the contract allows accounting changes to affect contract calculations and contains an accounting-based performance-pricing provision but not an accounting-based dividend restriction, 0 otherwise;

Div_Res = 1 if the contract allows accounting changes to affect contract calculations and contains an accounting-based dividend restriction but not an accounting-based performance-pricing provision, 0 otherwise; and

PP&Div_Res = 1 if the contract allows accounting changes to affect contract calculations and contains both an accounting-based performance pricing provision and an accounting-based dividend restriction, 0 otherwise.

All other variables are as defined in Model (1).

Since all borrowers in our sample have some type of accounting-based covenant, this research design allows us to examine the incremental effect of performance pricing and to differentiate between covenants that are linked to dividend restrictions vs. those that are not.

IV. SAMPLE

To obtain a sample of firms that have bank debt and also made material voluntary accounting changes, we conduct a keyword search on the Lexis/Nexis database.[3] When a firm makes a nonmandated accounting change, the Securities and Exchange Commission

---

[3] We searched the 10-Ks and 10-Qs filed on the Lexis/Nexis database using the phrase "exhibit 18" and "Letter re Change in Accounting Principles."

(SEC) requires the firm's independent accountant to issue a "preferability letter" commenting on whether they prefer the new accounting method relative to the prior method. The firm must file these letters with the SEC as exhibit 18 in their annual 10-K or quarterly 10-Q financial statements.

To identify firms that change accounting methods, we search all 10-K and 10-Q reports available on Lexis/Nexis from 1995 through 2000 for exhibit 18 disclosures. Then we read each exhibit 18 report to determine whether the firm made a material accounting change. Table 1 describes the results of this sample selection process. This search yields a sample of 296 firms that made accounting changes from January 1, 1995, through June 30, 2000. From these 296 firms, we exclude 48 that made accounting changes that had no effect on income statement or balance sheet accounts (e.g., changes in the format of the income statement, changes in the definition of a cash equivalent, or immaterial changes in accounting methods) and we also exclude two firms whose preferability letters were actually for mandatory accounting changes. We also excluded five firms that were banks or had banks as significant subsidiaries, because banks' accounting method choices may be affected by regulatory incentives.

We then used the debt footnotes to determine the amount and type of debt outstanding, and used the exhibit list at the end of each filing to determine whether the borrower filed a debt contract and, if so, the exact filing that includes that debt contract. Based on this

TABLE 1
Sample Selection Process
Borrowers Making Material Voluntary Accounting Method Changes during 1995–2000

Firms that filed "preferability letters" with the SEC[a]	296
Less:	
Firms that made accounting changes that had no effect on balance sheet or income statement accounts[b]	48
Financial institutions	5
Firms that made mandatory accounting changes improperly classified as voluntary changes	2
Firms with no bank debt or with immaterial bank debt	65
Nonfinancial firms with material bank debt that made material voluntary accounting changes	176
Less:	
Firms with bank debt that exceeds the SEC's materiality threshold of 10 percent of assets, but do not report filing the debt contract with the SEC in their 10-K exhibit list	24
Firms that indicated in their 10-K exhibit list that bank debt was filed with the SEC, but we could not find the contract on EDGAR[c]	27
Firms in our sample	125

---

[a] When a firm makes a nonmandated accounting change, the SEC requires the firm's independent accountant to issue a "preferability letter" commenting on the preferability of the new accounting method relative to the prior method.

[b] Examples include changes in the format of the income statement, changes in the definition of a cash equivalent, or immaterial changes in accounting methods.

[c] Although most of the indicated filings are available on global access, we cannot confirm whether the debt contracts were filed or examine filed contracts because our subscription does not allow us to view most of these filings.

analysis, we excluded another 65 firms with no bank debt or with immaterial bank debt.[4] These adjustments left a sample of 176 nonfinancial firms with material bank debt that also made voluntary material changes in accounting methods.

For each of these 176 borrowers, we systematically searched their debt footnotes, 10-Ks, 10-Qs, 8-Ks, and all their registration statements filed online to obtain copies of their debt contracts. We identified 24 firms whose footnotes indicate that they have a material bank debt contract, but whose exhibit indexes indicate that they did not file copies of the bank debt with the SEC. More specifically, we examine the exhibit index in the borrower's 10-K filed before and after the accounting change to determine whether the borrower filed the bank debt contract. For each of these 24 firms, no reference to any bank debt contracts appeared in the entire exhibit index. If the exhibit indexes of the 10-Ks filed before and after the accounting change did not mention bank debt, we assume that the borrower never filed it with the SEC.[5] In addition, we could not access debt contracts for another 27 firms, even though their exhibit index indicates that they filed bank debt contracts with the SEC.[6] Because our study focuses on the effects of debt contract provisions such as covenants, dividend restrictions, and performance pricing on borrowers' voluntary accounting changes, and we do not know whether these debt contracts allowed accounting changes to affect contract calculations, we must exclude these 51 borrowers from the sample.

This sample selection process yielded our final sample of 125 firms.[7] Although our sample attrition rate is high, we argue that our inferences are unlikely to be an artifact of sample selection bias for two reasons. First, we examine the types of accounting changes made by the firms we exclude, and the effects of these changes on net income, to ensure that our data requirements are not systematically related to our dependent variable. We find these firms' accounting changes are remarkably similar to those made by the sample firms. Of the 51 firms excluded from the sample, 29 (57 percent) made income-increasing accounting changes and 22 (43 percent) made income-decreasing accounting changes, compared to 54 percent increases and 46 percent decreases in our sample firms. The difference between these two groups is not statistically significant. As for the sample firms, the most common accounting changes for the excluded firms were, in order of frequency, for inventory, depreciation/capitalization, and revenue recognition. Thus, we conclude that our data requirements did not result in an endogenous sample selection bias.

Second, we compare other characteristics of our sample firms and the excluded firms to ensure that these two groups of firms are not systematically different on other dimensions. Specifically, we compare asset sizes, leverage ratios, and the proportion of borrowers reporting small and large losses before the effect of the accounting changes. None of these characteristics was significantly different between our sample and the excluded borrowers. Furthermore, we confirmed that these characteristics did not differ between the sample and excluded borrowers after partitioning on the accounting changes' effect on net income.

---

[4] Rule 601 of regulation SK allows borrowers some latitude in filing their debt contracts. Specifically, the SEC requires borrowers to file a copy of a debt contract if it is material, where the SEC generally defines materiality as 10 percent of the firm's assets. See Press and Weintrop (1990) for more discussion of the SEC's filing requirements for private debt contracts. We classified firms as having immaterial bank debt if no bank debt filings were listed in their exhibit indexes and the amount of the bank debt disclosed in their footnotes was less than 10 percent of their assets.

[5] We tested the validity of this assumption by conducting exhaustive online searches for debt contracts for each of these firms, and we were unable to locate copies of any of the bank debt outstanding at the time the firm made the accounting change. The inability to find any contracts online supports the assumption that the 24 firms never filed these debt contracts with the SEC.

[6] Although most of the indicated filings are available on the Global Access database, we can neither confirm whether these 27 firms filed nor can we examine any contracts these 27 firms might have filed because our subscription to Thomson Financial does not allow us to view most of these filings.

[7] Our logit regressions require additional data, reducing the sample size to 121 in that analysis.

Descriptive Statistics

Table 2 provides descriptive evidence on our sample firms' accounting method changes. The sample is split fairly evenly between income-increasing and income-decreasing accounting changes, with 67 increases and 58 decreases. The most common change is in the method of accounting for inventory, which represents 42 percent of our sample. Of the 53 inventory accounting changes, 41 are changes from LIFO to some other inventory accounting method, 6 are changes from FIFO to LIFO, and 6 are other changes in inventory methods, such as changes in overhead allocation. Changes in capitalization of assets, depreciation, and revenue recognition methods each represent roughly 15 percent of the voluntary accounting changes in our sample. Nearly two thirds of the changes in capitalization reduced income, whereas the split between increases and decreases is fairly even for depreciation changes and changes in revenue recognition as well as for the sample as a whole.

V. RESULTS

Univariate Analysis

Table 3 presents univariate analyses of the direction and (signed) magnitude (scaled by assets before the effect of the change) of the income effect of the accounting method change. Panel A of Table 3 presents the actual and expected (assuming independence) number of income-increasing vs. income-decreasing accounting changes for borrowers whose debt contracts allow voluntary accounting changes to affect calculations vs. those that do not. We find that 75 borrowers have at least one debt contract that allows voluntary changes to affect contract calculations, and 50 of our borrowers do not have any bank debt contracts that allow voluntary accounting changes to affect calculations. The accounting

TABLE 2
Accounting Method Changes Partitioned by Effect on Income for Borrowers Making Voluntary Accounting Method Changes during 1995–2000

Type of Accounting Change	*Total*	*Income-Increasing*	*Income-Decreasing*
Capitalizing to expensing costs or expensing to capitalizing	19	7	12
Depreciation policies	18	10	8
Gains on pension assets	4	4	0
Goodwill	5	0	5
Inventory accounting—from LIFO	41	23	18
Inventory accounting—from FIFO	6	4	2
Inventory accounting—Other	6	6	0
Revenue recognition	14	6	8
Other	12	7	5
Total	125	67	58

TABLE 3

Tests of Independence between Whether Accounting Changes Affect Contract Calculations and the Sign and Magnitude of the Effect of the Accounting Change on Income

Panel A: Contingency Table of Actual and (Expected) Number of Firms Making Accounting Changes, Testing for Independence between Whether Accounting Changes Affect Contract Calculations and Whether Accounting Change is Income-Increasing vs. Income-Decreasing

	Accounting Method Changes Affect Contract Calculations	*Accounting Method Changes Do Not Affect Contract Calculations*	*Total*
Income-increasing accounting change	46 (40.2)	21 (26.8)	67
Income-decreasing accounting change	29 (34.8)	29 (23.2)	58
Total	75	50	125
Test of independence	Chi-square 4.51	p-value 0.04	

Panel B: Wilcoxon Rank Sum Test of Independence of Distribution of Income Effect of Accounting Method Change for Borrowers with Debt Contract Calculations Affected or Not Affected by Accounting Method Changes

	Accounting Method Changes Affect Contract Calculations	*Accounting Method Changes Do Not Affect Contract Calculations*
Actual rank sum	69.42	53.36
Expected rank sum	(63)	(63)
Test of independence	Z-statistic 2.42	p-value 0.008

method change increases income for 61 percent (46/75) of borrowers whose debt contracts allow the changes to affect calculations vs. only 42 percent (21/50) of borrowers whose contracts do not allow voluntary accounting changes to affect the calculations. The Chi-square statistic for independence of these two characteristics is 4.51, which is significant at the 4 percent level.

Panel B of Table 3 presents the actual and expected (assuming independence) rank sum of the magnitude of the income effect of the accounting method change for borrowers whose debt contracts allow voluntary accounting changes to affect calculations vs. those that do not. We find that when accounting method changes affect the contract calculations, borrowers make accounting changes that more positively affect income. We perform a Wilcoxon rank sum test of the distribution of the income effect of the accounting change for borrowers with debt contract calculations affected by accounting method changes vs. those without, and reject independence of the distribution of the income effect of the accounting change across these two types of borrowers at the 0.008 level.

Table 4 presents the means and standard deviations of our independent variables, partitioned by the direction in which the accounting change affects income, as well as t-statistics for the differences between the means of these two groups. Several of the debt-contracting variables differ significantly between borrowers making income-increasing (vs.

TABLE 4
Mean and (Standard Deviation) of Independent Variables Partitioned by the Effect of the Accounting Method Change on Income

Variable	*Predicted Sign of Difference*	*Income-Increasing Mean (std. dev.)*	*Income-Decreasing Mean (std. dev.)*	*Difference Mean (t-statistic)*
Debt Contracting Variables				
Acc_Change	+	0.69 (0.46)	0.50 (0.50)	0.19 (2.15)**
NoPPorDiv_Res	+	0.09 (0.29)	0.19 (0.40)	−0.10 (−1.63)
PP	+	0.13 (0.34)	0.07 (0.26)	0.06 (1.22)
Div_Res	+	0.25 (0.44)	0.14 (0.44)	0.11 (1.65)*
PP&Div_Res	+	0.21 (0.41)	0.10 (0.31)	0.11 (1.64)*
One_Lender	±	0.09 (0.29)	0.17 (0.38)	−0.08 (−1.38)
Control Variables				
COMP	+	0.57 (0.50)	0.34 (0.48)	0.23 (2.53)***
SMLOSS	+	0.18 (0.39)	0.16 (0.37)	0.02 (0.35)
SMLOSS_NL	±	0.03 (0.39)	0.12 (0.37)	−0.09 (−1.89)*
LGLOSS	−	0.19 (0.40)	0.31 (0.47)	−0.12 −(1.63)*
LGLOSS_ΔCEO	−	0.02 (0.40)	0.12 (0.47)	−0.10 (−2.41)***
NOL	+	0.12 (0.33)	0.16 (0.37)	−0.04 (−0.58)
SIZE	±	13.51 (1.89)	13.18 (1.83)	0.33 (1.00)
NLIST	−	0.15 (0.36)	0.36 (0.48)	−0.21 (−2.81)***
ΔCEO	−	0.08 (0.27)	0.18 (0.38)	−0.10 (−1.59)*
Number of Firms[a]		67	58	

(continued on next page)

TABLE 4 (continued)

***, **, and * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively, for one- or two-tailed tests as appropriate.

[a] We could not obtain proxy statements for four borrowers, so there are only 121 observations for the COMP, LGLOSS_ΔCEO, and ΔCEO variables.

Variable Definitions

Debt Contracting Variables:

ACC_Change = 1 if contract allows accounting changes to affect calculations, 0 otherwise;

NoPPorDiv_Res = 1 if the contract allows accounting changes to affect calculations but contains neither an accounting-based performance pricing provision nor an accounting-based dividend restriction, 0 otherwise;

PP = 1 if the contract allows accounting changes to affect calculations and contains an accounting-based performance pricing provision but not an accounting-based dividend restriction, 0 otherwise;

Div_Res = 1 if the contract allows accounting changes to affect calculations and contains an accounting-based dividend but not an accounting-based performance pricing provision, 0 otherwise;

PP&Div_Res = 1 if the contract allows accounting changes to affect calculations and contains both an accounting-based performance pricing provision and an accounting-based dividend restriction, 0 otherwise; and

One_lender = 1 if accounting method changes affect debt contract calculations *and* a single lender makes the loan, 0 otherwise.

Control Variables:

COMP = 1 if the firm indicates that the managers' bonuses depend on financial performance *and* the CEO received a bonus in the year prior to the accounting change, 0 otherwise;

SMLOSS = 1 if the firm's earnings, before the effect of the accounting change, is between 0 and −1 percent of assets, 0 otherwise;

SMLOSS_NL = 1 if the firm's earnings, before the effect of the accounting change, is between 0 and −1 percent of assets *and* the firm is not listed on a major stock exchange, 0 otherwise;

LGLOSS = 1 if the firm's earnings, before the effect of the accounting change, is less than −1 percent of assets, 0 otherwise;

LGLOSS_ΔCEO = 1 if the firm's earnings, before the effect of the accounting change, is less than −1 percent of assets *and* the firm's CEO changed in the year prior to the accounting change, 0 otherwise;

NOL = 1 for borrowers that have a net operating loss carryforward *and* make an accounting change that has a tax effect, 0 otherwise;

SIZE = the log of the firms assets before the effect of the accounting change;

NLIST = 1 for firms not listed on a major stock exchange, 0 otherwise; and

ΔCEO = 1 if the firm's CEO changed in the year prior to the accounting change, 0 otherwise.

income-decreasing) accounting changes. Voluntary changes in accounting methods affect debt contract calculations for 69 percent of the firms that make income-increasing accounting method changes, compared to only 50 percent of those that make income-decreasing changes. Borrowers that make income-increasing accounting method changes are more likely to face dividend restrictions affected by those changes than are borrowers who make income-decreasing accounting changes.

Table 4 also reports that some of our control variables differ significantly between borrowers who make income-increasing vs. income-decreasing changes. Borrowers who pay their executives bonuses based on accounting performance are more likely to make income-increasing accounting method changes. An income-increasing accounting method change is less likely to be made by an unlisted firm that reports a small loss, or by a borrower that reports a large loss, especially if the firm changed its CEO in the year of the accounting change.

Correlations among Independent Variables

Fields et al. (2001) suggest that a limitation of previous accounting choice studies is that they typically examine only one potential motive for making accounting choices, without controlling for other possible motives. The correlations among our independent variables reported in Table 5 provide evidence on the extent of the "correlated omitted variable" problem associated with focusing on a single motive.

In general, we observe no significant correlations between our debt-contracting variables (ACC_Change, NoPPorDiv_Res, PP, Div_Res, and One_Lender) and other factors that might influence the accounting choice decision (COMP, SMLOSS, SMLOSS_NL, LGLOSS, LGLOSS_ΔCEO, NOL, SIZE, NLIST, and ΔCEO). In particular, whether a debt contract allows accounting method changes to affect contract calculations (ACC_Change) is not significantly correlated with any of the control variables. The second column of Table 5 does show, however, that when accounting changes affect contract calculations and no accounting-based performance-pricing or dividend restrictions exist (i.e., NoPPorDiv_Res =1), borrowers are less likely to pay their executives a bonus based on accounting performance and are more likely to report large losses. These correlations suggest that it may be important to control for management compensation and incentives to meet earnings benchmarks when we test our debt-contracting hypotheses.

Many of our independent variables are correlated with borrower size. Specifically, larger borrowers are more likely to have accounting-based dividend restrictions that are not combined with performance pricing, and are more likely to pay their executives accounting-based performance bonuses. Meanwhile, larger borrowers' debt contracts are less likely to allow accounting changes to affect contract calculations when there are no dividend restrictions or performance pricing. Larger borrowers are also less likely to report a large loss before the effect of the accounting change, or an NOL, and are less likely not to list their shares on a major exchange. Not surprisingly, borrowers with NOLs are more likely to report losses, either large or small; are more likely to experience CEO turnover; and are less likely to pay their executives a bonus based on accounting performance.

Regression Analysis

Table 6 reports the results of our logistic regressions. Model (1), which we use to test Hypotheses 1 and 2, has explanatory power of 24.3 percent. This model correctly classifies 81.6 percent of the accounting method changes as increasing or decreasing income, compared to a naïve prediction accuracy rate of 55.4 percent if we assume that all accounting changes increase income. The second model, which we use to test Hypotheses 3 and 4, has slightly higher explanatory power at 27.8 percent and correctly classifies 84 percent of the accounting changes as either income-increasing or income-decreasing.[8]

The significantly positive coefficient on ACC_Change in Model (1) indicates that borrowers that change accounting methods are more likely to make changes that increase income when the changes affect debt contract calculations. Following Amemiya (1981, 1488), we convert the estimated coefficient to a change in probability by multiplying the estimated logit coefficient by 0.25. Borrowers are 39 percent (1.57 × 0.25) more likely to make income-increasing (rather than income-decreasing) accounting changes when their

---

[8] The sign and magnitude of the effect of the accounting change are the same on both net income and net worth when the firm accounts for changes in accounting methods using either the cumulative effect or prospective method. However, 22 of our sample borrowers use the retroactive approach, where changes in accounting methods have different income statement and net worth effects. Excluding these firms from our analyses does not alter our conclusions.

TABLE 5
Pearson Correlations among Independent Variables
(significance levels in parentheses)

	ACC_ Change	*NoPPor Div_Res*	*PP*	*Div_ Res*	*PP&Div_ Res*	*One_ Lender*	*COMP*	*SM LOSS*	*SM LOSS_NL*	*LG LOSS*	*LGLOSS_ ΔCEO*	*NOL*	*SIZE*	*NLIST*
ACC_Change	1.00													
NoPPorDiv_Res	0.33 (0.00)	1.00												
PP	0.28 (0.00)	−0.14 (0.13)	1.00											
Div_Res	0.41 (0.00)	−0.19 (0.02)	−0.17 (0.06)	1.00										
PP&Div_Res	0.36 (0.00)	−0.17 (0.05)	−0.15 (0.10)	−0.22 (0.01)	1.00									
One_Lender	0.31 (0.00)	0.27 (0.00)	−0.13 (0.15)	0.23 (0.00)	0.03 (0.76)	1.00								
COMP	0.01 (0.94)	−0.23 (0.01)	−0.00 (0.99)	0.12 (0.19)	0.12 (0.19)	0.03 (0.76)	1.00							
SMLOSS	0.02 (0.80)	0.01 (0.92)	−0.01 (0.89)	−0.06 (0.48)	0.10 (0.29)	−0.04 (0.63)	0.05 (0.55)	1.00						
SMLOSS_NL	0.04 (0.67)	−0.02 (0.82)	0.11 (0.23)	−0.06 (0.49)	0.05 (0.60)	−0.01 (0.88)	−0.01 (0.90)	0.62 (0.00)	1.00					
LGLOSS	0.09 (0.31)	0.20 (0.02)	0.11 (0.23)	−0.15 (0.10)	0.00 (0.98)	0.06 (0.53)	−0.09 (0.33)	−0.26 (0.01)	−0.16 (0.07)	1.00				
LGLOSS_ΔCEO	0.09 (0.34)	0.19 (0.04)	0.12 (0.18)	−0.13 (0.15)	−0.02 (0.85)	0.10 (0.27)	−0.19 (0.04)	−0.12 (0.20)	−0.07 (0.44)	0.46 (0.00)	1.00			

(*continued on next page*)

TABLE 5 (continued)

	ACC_ Change	*NoPPor Div_Res*	*PP*	*Div_ Res*	*PP&Div_ Res*	*One_ Lender*	*COMP*	*SM LOSS*	*SM LOSS_NL*	*LG LOSS*	*LGLOSS_ ΔCEO*	*NOL*	*SIZE*	*NLIST*
NOL	0.04 (0.67)	0.05 (0.60)	0.09 (0.30)	−0.08 (0.37)	0.02 (0.84)	−0.01 (0.89)	−0.18 (0.04)	0.00 (0.92)	0.16 (0.07)	0.31 (0.00)	0.27 (0.00)	1.00		
SIZE	−0.04 (0.67)	−0.24 (0.01)	−0.07 (0.43)	0.24 (0.01)	−0.04 (0.70)	−0.39 (0.00)	0.21 (0.02)	0.10 (0.29)	−0.10 (0.27)	−0.22 (0.01)	−0.20 (0.02)	−0.16 (0.08)	1.00	
NLIST	−0.02 (0.80)	0.04 (0.64)	−0.01 (0.88)	−0.10 (0.26)	0.05 (0.56)	0.06 (0.53)	−0.16 (0.07)	0.19 (0.04)	0.49 (0.00)	0.14 (0.11)	0.16 (0.08)	0.10 (0.28)	−0.34 (0.00)	1.00
ΔCEO	0.06 (0.51)	0.08 (0.41)	0.11 (0.22)	0.00 (0.98)	−0.09 (0.34)	0.16 (0.07)	−0.11 (0.72)	−0.03 (0.72)	0.10 (0.27)	0.25 (0.01)	0.70 (0.00)	0.28 (0.00)	−0.17 (0.35)	0.20 (0.03)

Variables are defined in Table 4.

TABLE 6
Coefficients and (t-statistics) from Logit Regressions of the Sign of the Income Effect of Voluntary Accounting Changes on Debt Contracting and Control Variables
Sample of 64 Income-Increasing and 57 Income-Decreasing Voluntary Accounting Method Changes, 1995–2000

Variable	*Predicted Sign*	*Model (1) Coefficient (t-statistic)*	*Model (2) Coefficient (t-statistic)*
Intercept	+	2.36 (1.28)	2.92 (1.49)
Debt Contracting Variables			
Acc_Change	+	1.57 (3.19)***	
NoPPorDiv_Res	+		0.20 (0.26)
PP	+		2.34 (2.42)***
Div_Res	+		1.84 (2.72)***
PP&Div_Res	+		2.01 (2.52)***
One_lender	−	−2.36 (−2.99)***	−2.41 (−2.72)***
Control Variables			
COMP	+	1.31 (2.71)***	1.15 (2.32)***
SMLOSS	+	1.69 (1.87)**	1.96 (2.14)**
SMLOSS_NL	−	−5.13 (−3.23)***	−5.91 (−3.42)***
LGLOSS	−	−1.01 (−1.69)**	−0.96 (−1.55)*
LGLOSS_ΔCEO	−	−2.04 (−1.62)*	−2.59 (−1.78)**
NOL	+	1.14 (1.49)*	1.29 (1.62)*
SIZE	±	0.24 (1.74)*	0.28 (1.89)*
Pseudo R^2		24.3%	27.8%
Percent Correctly Predicted		81.6	84.0

***, **, and * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively, for one- or two-tailed tests as appropriate.

Variables are defined in Table 4.

debt contracts allow the accounting changes to affect contract calculations.[9] This supports our first hypothesis, that debt contracting is an important consideration in borrowers' accounting method change decisions.

The significant negative coefficient on One_Lender is consistent with our second hypothesis that borrowers that change accounting methods are less likely to make income-increasing accounting changes when the expected costs of technical default are lower, as should be the case when all of their bank debt contracts are from a single lender.

Model (2) partitions ACC_Change into four mutually exclusive subgroups to provide evidence on our third hypothesis (that performance pricing provides additional incentives beyond those provided by traditional covenants) and our fourth hypothesis (that the incentives provided by covenants restricting dividend payments, either directly or indirectly, may differ from those provided by other covenants). The significantly positive coefficient on PP indicates that borrowers that make accounting changes are 59 percent (2.34 × 0.25) more likely to make income-increasing accounting method changes when the changes not only affect their covenants, but also affect their interest rates through performance-pricing provisions. This supports our third hypothesis. The coefficient on Div_Res indicates that borrowers are 46 percent (1.84 × 0.25) more likely to make income-increasing (rather than income-decreasing) accounting method changes when the changes affect their dividend restrictions. This is consistent with our fourth hypothesis. The coefficient on PP&Div_Res indicates that borrowers with both performance-pricing and dividend restrictions are 50 percent (2.01 × 0.25) more likely to report income-increasing accounting method changes than are borrowers with neither performance-pricing nor dividend restrictions. However, the coefficient on this variable is not significantly different from the coefficients on PP or on Div_Res. The insignificant coefficient on NoPPorDiv_Res indicates that we find no evidence that borrowers without performance-pricing or dividend restrictions were more likely to report income-increasing accounting method changes.

We also find some evidence that management compensation incentives influence the choice between income-increasing and income-decreasing accounting method changes. Specifically, when managers receive accounting-based performance bonuses (COMP) they are more likely to make income-increasing changes. In addition, consistent with a "big bath" story, borrowers with new CEOs who report large losses before the effects of the accounting method changes (LGLOSS_ΔCEO) are more likely to report income-decreasing accounting changes than income-increasing changes.

We also find some evidence that noncontracting motives to manage earnings affect borrowers' decisions to make income-increasing vs. income-decreasing accounting method changes. Specifically, we find borrowers with shares listed on a major exchange expecting to report a small loss prior to the accounting method change (SMLOSS) are more likely to report an income-increasing change, while those not listed on a major exchange are more likely to report an income-decreasing change (SMLOSS_NL). This is consistent with the Beatty, Ke, and Petroni (2002) conclusion that exchange-listed firms are more likely to manage earnings to beat earnings benchmarks than are firms that are not listed on a major exchange.

Finally, we find some evidence that external parties affect accounting choices. Specifically, borrowers with NOL carryforwards who make accounting changes affecting taxes are more likely to make income-increasing accounting changes than borrowers without such carryforwards.

---

[9] The conversion factor is obtained by taking the derivative of the predicted probability with respect to a particular independent variable. For a linear probability model that derivative equals the estimated coefficient. For a logit model that derivative equals the estimated coefficient multiplied by the exponential of the predicted probability divided by the square of 1 plus the exponential of the predicted probability.

VI. SENSITIVITY ANALYSIS

Signed Magnitude of Income Effect

The tests of our hypotheses reported in Table 6 examine the sign of the income effect of the accounting method change but not the magnitude of the effect. In Table 7, we report the results of rank regressions of the signed magnitude of the income effect of the accounting change, scaled by assets before the effect of the change. The results are largely consistent with those from our logistic analyses. As our first hypothesis predicts, borrowers make accounting changes that have more positive effects on income when the accounting changes affect their debt contract calculations. Consistent with our second hypothesis, borrowers make accounting changes that have a less positive effect on income when the accounting changes affect their debt contract calculation, but the costs of covenant violation are lower because all of their bank debt contracts are with a single lender. With respect to our third and fourth hypotheses, borrowers whose bank debt has performance-pricing and dividend restrictions make accounting changes that have a more positive effect on income and, in fact, these appear to be the two most significant debt-contracting factors explaining the magnitude of the income effect. The results on our control variables largely conform to the results from our previous analyses. However, the performance-based executive bonus compensation variable (COMP) is insignificant in our rank regressions.

We examine the sensitivity of our rank regression results to two alternative specifications. First, we re-estimate the regressions without ranking the data. After winsorizing the dependent variable for three outliers that fall more than 5 standard deviations from the mean, we obtain results similar to those in Table 7. We also obtain similar results when we re-estimate the rank regressions using the unscaled income effect of the accounting method change as the dependent variable.

Debt-Contracting Variables

We consider all bank debt contracts when we construct our debt-contracting variables. However, these variables do not incorporate covenants on non-bank debt. Since covenants on non-bank debt typically allow accounting changes to affect contract calculations (see Smith and Warner 1979), we include a dichotomous variable (PUB_COV) equal to 1 for firms with covenants on non-bank debt. When we include PUB_COV as another debt-contracting variable its coefficient is insignificant and our other results remain similar. We also include another variable equal to 1 if the borrower has public debt and accounting changes do not affect bank-debt calculations, 0 otherwise (PUB_COV × [1 − ACC_COV]). Again, the coefficient on this variable is insignificant, indicating that covenants on public debt do not appear to affect whether borrowers' accounting changes are income-increasing vs. income-decreasing, regardless of whether these accounting changes affect bank debt calculations. Our other results remain similar.

Our debt-contracting variables capture the average effect of accounting-based covenants on the decision to record an income-increasing vs. an income-decreasing accounting change. However, this relation is likely to differ cross-sectionally, reflecting differences in expected cost of future violations. Since our One_Lender variable measures these differences only indirectly, we considered an alternative (indirect) proxy as an additional debt-contracting variable in our analyses. Following El-Gazaar and Pastena (1991), we consider the number of covenants as a measure of tightness of covenant restrictions. The number of covenants is not significant in our models. This may indicate either that the initial covenant tightness

TABLE 7
Coefficients and (t-statistics) from Rank Regressions of the Effect of Voluntary Accounting Changes on the Income Effect of the Accounting Change (Scaled by Assets Before the Effect of the Change) on Contracting and Control Variables
Sample of 64 Income-Increasing and 57 Income-Decreasing Voluntary Accounting Method Changes, 1995–2000

Variable	*Predicted Sign*	*Model (1) Coefficient (t-statistic)*	*Model (2) Coefficient (t-statistic)*
Intercept	+	50.30 (5.79)***	53.14 (5.99)***
Debt Contracting Variable			
ACC_Chang	+	20.98 (3.21)**	—
NoPPorDiv_Res	+	—	4.57 (0.44)
PP	+	—	34.25 (3.27)***
Div_Res	+	—	26.25 (2.86)***
PP&Div_Res	+	—	15.82 (1.70)**
One_Lender	−	−32.27 (−3.04)**	−29.57 (−2.63)***
Control Variables			
COMP	+	8.58 (1.34)*	6.65 (1.02)
SMLOSS	+	27.56 (2.62)***	32.02 (3.04)***
SMLOSS_NL	−	−41.56 (−2.59)**	−47.60 (−2.97)***
LGLOSS	−	−0.95 (0.11)	0.64 (0.08)
LGLOSS_ΔCEO	−	−20.53 (−1.45)*	−20.14 (−1.43)*
NOL	+	17.16 (1.76)**	16.42 (1.71)**
SIZE	±	−0.05 (−0.53)	−0.09 (−0.87)
Adjusted R^2		15.25%	17.8%

***, **, and * Indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively, for one- or two-tailed tests as appropriate.
Variable are defined in Table 4.

is not an important determinant of the subsequent accounting choice decision, or that the number of covenants is not a good proxy for covenant tightness.

VII. CONCLUSIONS

Fields et al. (2001, 301) argue that "research on accounting choice addresses the fundamental question of whether accounting matters." However, they also state that "the evidence on whether accounting choices are motivated by debt covenant concerns is inconclusive." By analyzing contract-specific details, we provide sharper evidence on the importance of debt contracts in borrowers' accounting choices. Specifically, our tests capitalize on the facts that (1) many recent debt contracts do not allow borrowers to use voluntary accounting method changes to affect contract calculations, (2) debt contracts often include performance pricing features as well as debt covenants, and (3) debt contracting decisions may affect accounting choices of borrowers who are not approaching covenant violations.

We find that borrowers that make accounting method changes are more likely to make income-increasing changes when their debt contracts allow these changes to affect contract calculations. This increase in likelihood of an income-increasing accounting change is lower when the borrower expects costs of technical violation to be lower because all of the borrower's bank debt is from a single lender. These results, which hold even after controlling for other motives for changing accounting methods (such as executive compensation incentives, incentives to meet earnings benchmarks, and incentives to reduce taxes), suggest that debt contracts influence borrowers' accounting choices. We also find that borrowers that voluntarily change their accounting methods are more likely to make income-increasing changes if their debt contracts include accounting-based performance-pricing or dividend restrictions. These results suggest that incentives to lower interest rates through performance pricing or to retain dividend payment flexibility influence borrowers' accounting method choices, thereby providing indirect evidence addressing the Fields et al. (2001) fundamental questions of whether, under what circumstances, and how accounting choice matters.

Our focus on contract-specific details for borrowers who need not be approaching covenant violations creates two important data limitations. First, our tests do not include borrowers that do not file their bank debt contracts with the SEC because contract-specific information is unavailable for these borrowers. This could affect the generalizability of our results if the relation between debt contracting and accounting choice differs for these borrowers. To mitigate this concern we provide evidence that these borrowers made voluntary accounting changes that are similar to those made by our sample borrowers. In addition, we document that the excluded borrowers and sample borrowers are similar along a number of dimensions including firm size, income before the accounting change, and leverage.

Second, we cannot directly test whether an accounting change (1) reduced the interest rate charged on the loan based on the performance pricing grid, (2) increased the borrower's flexibility to pay dividends, or (3) reduced the likelihood the borrower would violate a covenant. Since our sample includes borrowers that may be making accounting changes not only to affect the current debt contracting calculations, but also future calculations, we would need to estimate the effect of their accounting change on the interest rate, dividends paid, and covenant violations both in the quarter of the accounting change, and in all future quarters that could be affected by the accounting change as well. Since most borrowers have equity-based covenants and dividend restrictions, this means an accounting change would affect all future calculations until the debt matures. In addition, it is difficult to develop such estimates using only publicly available current period data. For example, when we use balance sheet data available in the financial statements, the ratios we calculate are not equal to the ratios used in the debt contracts. For all these reasons, we leave it to future research to quantify the effects of borrowers' voluntary changes in accounting methods on their interest rates, dividend payments, and probability of covenant violations.

REFERENCES

Amemiya, T. 1981. Qualitative response model: A survey. *Journal of Economic Literature* 19 (December): 1483–1586.

Asquith, P., A. Beatty, and J. Weber. 2002. Performance pricing in private debt contracts. Working paper, Massachusetts Institute of Technology, Cambridge, MA.

Beatty, A., I. Dichev, and J. Weber. 2002. The role and characteristics of accounting-based performance pricing in private debt contracts. Working paper, University of Michigan, Ann Arbor, MI.

———, B. Ke, and K. Petroni. 2002. Earnings management to avoid earnings declines across publicly and privately held banks. *The Accounting Review* 77 (July): 547–570.

———, K. Ramesh, and J. Weber. 2002. The importance of accounting changes in debt contracts: The cost of flexibility in covenant calculations. *Journal of Accounting and Economics* 33 (June): 205–227.

Beneish, D., and E. Press. 1993. Costs of technical violation of accounting-based debt covenants. *The Accounting Review* 68 (April): 233–257.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* (December) 24: 99–126.

Dichev, I., and D. Skinner. 2002. Large-sample evidence on the debt covenant hypothesis. *Journal of Accounting Research* 40 (September).

El-Gazaar, S., and V. Pastena. 1991. Factors affecting the scope and initial tightness of covenant restrictions in private lending agreements. *Contemporary Accounting Research* 8 (Fall): 132–151.

Fields, T., T. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31 (September): 255–307.

Healy, P., and K. Palepu. 1990. Effectiveness of accounting-based dividend covenants. *Journal of Accounting and Economics* 12 (January): 97–124.

Mohrman, M. 1996. The use of fixed GAAP provisions in debt contracts. *Accounting Horizons* 10 (September): 78–91.

Press, E., and J. Weintrop. 1990. Accounting-based constraints in public and private debt agreements: Their association with leverage and impact on accounting choice. *Journal of Accounting and Economics* 12 (January): 65–95.

Smith, C., and J. Warner. 1979. On financial contracting: An analysis of bond covenants. *Journal of Financial Economics* 7 (June): 117–161.

Sweeney, A. 1994. Debt-covenant violations and managers' accounting responses. *Journal of Accounting and Economics* 17 (May): 281–308.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 143–168

Accounting and Stock Price Performance in Dynamic CEO Compensation Arrangements

John F. Boschen
College of William & Mary
Augustine Duru
American University
Lawrence A. Gordon
University of Maryland at College Park
Kimberly J. Smith
College of William & Mary

ABSTRACT: In this study we examine the long-run effects of unexpected firm performance on CEO compensation. We find that unexpectedly good accounting performance is initially associated with increases in CEO pay. However, this initial effect soon reverses, and is followed by lower CEO pay in later years. Overall, the CEO's long-run cumulative financial gain from unexpectedly good accounting performance is not significantly different from zero. In contrast, unexpectedly good stock price performance is associated with increases in CEO pay for several years. Thus, the CEO's long-run cumulative financial gain from unexpectedly good stock price performance is positive and significant.

Keywords: *executive compensation; stock price performance; accounting performance; dynamic models.*

Data Availability: *We include a list of sample firms in the paper. All other data are available from public sources identified in the text.*

We thank Bill Baber, John Core, Martin Loeb, Denise Jones, Larry Pulley, David Reeb, Aamer Sheikh, and Bob Thompson for useful suggestions. William Compton, Meredith Hines, Asif Dhala, Tom Pappalardo, Tracy Morse, and Melissa Smith provided excellent research assistance. We are particularly grateful to Linda Bamber (the editor), Raffi Indjejikian (the associate editor) and an anonymous reviewer for their substantial contributions to this paper.

Submitted November 1999
Accepted September 2002

I. INTRODUCTION

Prior research on the role that accounting measures of performance play in determining chief executive officer (CEO) compensation has yielded two primary conclusions. First, increases in accounting performance positively and significantly affect current CEO compensation, even after controlling for stock price performance (Antle and Smith 1986; Lambert and Larcker 1987; Jensen and Murphy 1990; Rosen 1992; Sloan 1993; Bushman and Smith 2001). Second, Antle and Smith (1986), Lambert and Larcker (1987), Rosen (1992), and Sloan (1993), show that accounting performance often has a greater effect on CEO compensation than does stock price performance.

We examine whether these conclusions continue to hold in the long run. Theory suggests that firms can mitigate agency problems by tying CEO pay to multiple periods of performance (Fama 1980; Lambert 1983; Rogerson 1985; Murphy 1986; Gibbons and Murphy 1992), and empirical evidence shows that CEOs do engage in long-term relationships with their firms (Warner et al. 1988; Weisbach 1988; Murphy and Zimmerman 1993; DeFond and Park 1999). Boschen and Smith (1995) show that the long-run effects of unexpected stock price performance on CEO compensation are substantially larger than the current-period effects. However, we currently know very little about how unexpected accounting performance affects CEO compensation in the long run.

Our results show that accounting performance and stock price performance have different long-run effects on CEO compensation. Like prior research, we find that unexpectedly good accounting performance is initially associated with higher CEO compensation. Surprisingly, however, we find that this initial increase is offset by *lower* pay in the future. Our results suggest that, over the long run, the CEO receives essentially no reward for an unexpected increase in accounting performance. In contrast, the CEO receives a substantial reward over the long run for an unexpected increase in stock price performance. Our evidence that the two main performance measures in compensation arrangements both have important multiperiod dynamics—particularly the puzzling result that better accounting performance is associated with lower future CEO compensation—suggests that future research on CEO incentives must address the multiperiod effects of performance on CEO pay arrangements.

Our analysis reveals that the traditional pay-performance model, based on first-differenced compensation, does not adequately capture the long-run effects of firm performance on CEO compensation. We show how researchers can adapt the vector autoregression (VAR) model that is widely used in the analysis of time series data (Hamilton 1994) to effectively capture these long-run relations. Finally, we conclude from our robustness tests that compensation models should allow for at least three lags of compensation and firm performance in order to fully capture the dynamics.

The paper proceeds as follows. In the next section, we review the findings of the prior literature. Section III describes the sample of firms we use, defines the variables, and presents descriptive statistics for the major variables. Section IV presents the empirical results that we obtain from estimating several versions of the dynamic pay-performance model. Section V summarizes and offers some concluding comments.

II. BACKGROUND AND PRIOR RESEARCH

Implied Dynamics of Traditional Compensation Models

Multiperiod agency models suggest that firms can mitigate both the moral hazard problem (Lambert 1983; Rogerson 1985; Murphy 1986) and the adverse selection problem (Fama 1980; Gibbons and Murphy 1992) by tying employee compensation to multiple years of performance. Relatively low turnover rates among chief executive officers (Warner et al. 1988; Weisbach 1988; Murphy and Zimmerman 1993; DeFond and Park 1999) indicate that chief executives do, in fact, have long-term relationships with their firms. However, empirical compensation researchers traditionally model changes in CEO compensation as a function of unexpected firm performance in the current period only (e.g., Lambert and Larcker 1987; Jensen and Murphy 1990; Sloan 1993).[1]

We present a simplified version of this traditional model (Lambert and Larcker 1987), with one performance measure, to show how compensation responds to unexpected firm performance in both the short run and the long run:

$$\Delta C_t = \beta_X[X_t - E(X_t)] + z_t^C, \tag{1}$$

where C_t and X_t, are compensation and firm performance in period t, E is the expectation operator, and z_t^C is the unexpected change in compensation. The response of compensation in period t to unexpected firm performance in period t (i.e., the contemporaneous response) is β_X.[2] Previous research has focused on this single parameter, β_X, as the measure of the pay-performance relation and has ignored the effects of unexpected firm performance on compensation over the long run. Although the model in Equation (1) does not explicitly estimate the long-run response of compensation to unexpected firm performance, the structure of the model implies that pay raises from unexpected firm performance remain *permanently* in the level of compensation. For example, if a CEO receives an extra \$100 in period t as a reward for his performance in period t, the model in Equation (1) predicts that he will also receive \$100 in period t + 1, another \$100 in period t + 2, and so on. Over a ten-year period, the CEO would receive a total of \$1,000 for his performance in period t.

To show why the model above implies that pay raises remain permanently in compensation, we rewrite Equation (1) in levels, with unexpected firm performance denoted as z_t^X (i.e., $[X_t - E(X_t)] = z_t^X$):

$$C_t = C_{t-1} + \beta_X z_t^X + z_t^C. \tag{2}$$

Then we write similar expressions for compensation in later periods, written in moving-average form, to illustrate the pattern that develops over time:

$$\begin{aligned}
C_t &= C_{t-1} + \beta_X z_t^X + z_t^C \\
C_{t+1} &= (C_{t-1} + \beta_X z_t^X + z_t^C) + \beta_X z_{t+1}^X + z_{t+1}^C \\
C_{t+2} &= (C_{t-1} + \beta_X z_t^X + z_t^C + \beta_X z_{t+1}^X + z_{t+1}^C) + \beta_X z_{t+2}^X + z_{t+2}^C \\
&\vdots \\
C_{t+k} &= (C_{t-1} + \beta_X z_t^X + z_t^C + \beta_X z_{t+1}^X + z_{t+1}^C + \beta_X z_{t+2}^X + z_{t+2}^C + \cdots + \beta_X z_{t+k-1}^X + z_{t+k-1}^C) \\
&\quad + \beta_X z_{t+k}^X + z_{t+k}^C \\
&\vdots
\end{aligned}$$

---

[1] Both Jensen and Murphy (1990) and Anderson et al. (2000) incorporate one lag of firm performance in their compensation equations, but they do not model the dynamic responses over time.

[2] As noted in Bushman and Smith (2001), early literature used the term *pay-performance sensitivity* to refer to the weight on a performance measure in a regression of compensation on performance—that is, the partial derivative of compensation with respect to firm performance (Murphy 1985). However, Banker and Datar (1989) use the term *sensitivity* to refer to the partial derivative of a performance measure with respect to an agent's action. To avoid ambiguity, and to be consistent with the time-series literature, we will use the term *response*, in the time-series sense, to denote the partial derivative of current (and future) compensation with respect to current unexpected firm performance.

The response of compensation in any period t + k, to unexpected firm performance in period t, is simply the partial derivative of C_{t+k} with respect to z_t^X. We say that the effect of unexpected firm performance on compensation is permanent when these derivatives are non-zero and identical for all periods. For the model above, this is clearly the case, since:

$$\partial C_t/\partial z_t^X = \partial C_{t+1}/\partial z_t^X = \partial C_{t+2}/\partial z_t^X = \cdots = \partial C_{t+k}/\partial z_t^X = \cdots = \beta_X.$$

The traditional compensation model in Equation (1) constrains the coefficient on prior compensation, C_{t-1}, to equal 1.0, as shown in Equation (2). However, if we relax this constraint, and allow the coefficient on C_{t-1} to be positive, but less than 1, then compensation responds to unexpected firm performance quite differently. We denote the coefficient on C_{t-1} as ρ and rewrite Equation (2) as;

$$C_t = \rho C_{t-1} + \beta_X z_t^X + z_t^C. \tag{2'}$$

We then follow the same process as above (i.e., we write the moving-average form for each C_{t+k} and take partial derivatives). However, now the responses of compensation to unexpected firm performance are no longer constant across periods. In fact, the derivatives showing the effects of unexpected firm performance on future CEO compensation (i.e., $\partial C_{t+k}/\partial z_t^X = \rho^k\beta_X$ for any k) approach zero as k becomes large. In this case, we say the effects of unexpected firm performance remain in compensation levels only *temporarily.* For example, if $\rho = 0.5$, then a CEO receiving an extra \$100 in period t as a reward for his performance in period t would receive only \$50 in period t + 1, \$25 in period t + 2, and so on, for a total of only \$200 over a ten-year period.

Boschen and Smith (1995) is the first study to model the long-run effects of firm performance (in their case, stock price performance) on CEO compensation. They show that, although the long-run response of CEO compensation to unexpected stock price performance is not permanent, it is long-lived, and the contemporaneous portion of the response is small relative to the long-run response. Thus, Boschen and Smith (1995) conclude that modeling only the contemporaneous portion of the compensation response misrepresents the long-run relation between CEO compensation and stock price performance. Their findings raise an important question about whether the contemporaneous relation between CEO compensation and accounting performance is also a misleading indicator of the long-run relation.

The Role of Accounting Performance in CEO Compensation

There are three non-mutually exclusive hypotheses as to why CEO compensation contracts include accounting performance in addition to stock price performance (Bushman and Smith 2001). These are: (1) the direct incentive hypothesis, in which accounting performance directly encourages work effort, (2) the filtering hypothesis, in which accounting performance filters out common noise in stock price performance, and (3) the effort allocation hypothesis, in which accounting performance shifts the CEO's work effort toward, or away from, a specific subset of actions. Below we discuss the implications of each of these hypotheses for the empirical relation between CEO pay and accounting performance.

The direct incentive hypothesis is based on the assumption that accounting performance provides another noisy signal about the CEO's effort, in addition to stock price performance. In this case, linking compensation to accounting performance allows the firm's owners to more closely tie the CEO's compensation to his work effort. Thus, this hypothesis implies a positive relation between accounting performance and CEO compensation. This positive relation does not require that accounting performance be less noisy than stock price performance. The essential requirement is that accounting performance contains independent information on CEO work effort.[3]

The filtering hypothesis is based on the assumption that the noise component in accounting performance is correlated with the noise component in stock price performance. In this case, linking compensation to accounting performance allows the firm's owners to partially or fully filter the noise from stock price performance. As with the direct incentive hypothesis, the goal is to more closely tie the CEO's compensation to his work effort (Banker and Datar 1989; Sloan 1993). This hypothesis predicts that the relation between CEO pay and accounting performance depends on the sign of the correlation between the noise components of accounting and stock price performance. For example, if the noise component in accounting performance is positively correlated with the noise component in stock price performance, then the filtering hypothesis implies a negative relation between accounting performance and CEO compensation.

The effort allocation hypothesis is based on the assumption that accounting performance only captures a subset of the CEO's actions, but stock price performance captures the effects of all CEO actions. In this case, linking compensation to accounting performance allows the firm's owners to provide the CEO with incentive to emphasize particular *types* of work effort (Bushman and Indjejikian 1993; Lambert 1993; Feltham and Xie 1994; Lambert 2001). For example, the firm's owners can provide an additional incentive for the subset of actions captured by accounting performance (above and beyond the incentive provided from their impact on stock price performance) by imposing a positive relation between accounting performance and CEO compensation. Alternatively, a negative relation between accounting performance and CEO compensation reduces the incentive for the subset of actions captured by accounting performance, by counteracting the incentive provided from the impact of these actions on stock price performance.

III. DATA

We use compensation and performance data for 30 firms over the 1959–1995 period, for a total of 1,110 firm-years. Appendix A lists the sample firms and their industries. Masson (1971) first presents a data set of compensation and performance data for 50 firms over the 1948–1966 period. Antle and Smith (1985, 1986) extend Masson's (1971) data set for the 39 firms still operating in 1977, and Boschen and Smith (1995) further extend it for the 16 firms remaining as of 1990. We extend Boschen and Smith's (1995) data set for the 15 firms still operating in 1995, and we also add data on 15 new firms, for which we are able to obtain full sets of proxy statements for the years 1959–1995. However, increasing the number of firms constrains our time period because we can obtain data on the new firms extending only back to 1959.

We use two measures of CEO compensation: (1) cash compensation, measured as the log of compensation paid in cash, and (2) total compensation, measured as the log of the sum of cash compensation, stock grants, stock option grants (valued using the Black and Scholes [1973] method), and any other noncash compensation. We use two measures of firm performance, accounting performance and stock price performance. We measure accounting performance as the annual rate of return on assets, calculated as earnings before extraordinary items divided by average total assets. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends. We adjust both performance measures for inflation and then standardize them (i.e., subtracting the mean and dividing by the standard deviation) to facilitate comparison of the pay-performance responses for stock price and accounting measures of performance.

---

[3] The amount of noise in accounting and stock price performance does affect the relative size of the coefficients on each measure in a compensation regression. Lambert and Larcker (1987) and Sloan (1993) show that when accounting performance has less noise (i.e., more precision) than stock performance, it has a larger impact on CEO compensation.

Table 1 presents descriptive statistics for our major variables. Panel A presents measures of central tendency and variation in the variables. The mean firm-specific average annual CEO cash and total compensation, expressed in 1982 dollars, are $830,000 and $1,150,000, respectively. We benchmark our compensation data for the last year of the sample period, 1995, by comparing our compensation data with compensation data for the Standard & Poor's (S&P) 500. Untabulated results show that our sample firms' average total CEO compensation does not differ significantly from the average total CEO compensation for the S&P 500. Furthermore, none of the main components of average compensation (i.e., salary, cash bonus, long-term incentive plans, restricted stock, or stock option grants) differs significantly between our sample firms and the S&P 500.

TABLE 1

Descriptive Statistics for CEO Compensation, Accounting Performance, and Stock Price Performance[a,b]

(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Central Tendency and Variation

	Mean	*Median*	*Standard Deviation*
Cash Compensation (in $1,000s)[c]	830	703	572
Total Compensation (in $1,000s)[c]	1,150	813	1,346
Accounting Performance	0.025	0.023	0.054
Stock Price Performance	0.102	0.077	0.303

Panel B: Simple Correlations[d]

	Cash Compensation	*Total Compensation*	*Accounting Performance*	*Stock Price Performance*
Cash Compensation	1.000	0.884	−0.097	0.073
Total Compensation		1.000	−0.120	0.100
Accounting Performance			1.000	0.095
Stock Price Performance				1.000

---

[a] We measure cash compensation as the sum of all compensation paid in cash. We measure total compensation as the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. We measure accounting performance as earnings before extraordinary items, scaled by average total assets. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends. We adjust all variables for inflation.

[b] We calculate the descriptive statistics in Panel A and simple correlations in Panel B with the data arranged in columns. The data for each firm is vertically stacked for firms 1 through 30.

[c] Cash and total compensation are expressed in 1982 dollars in Panel A.

[d] We use the natural log of cash and total compensation to calculate the correlations in Panel B.

Our accounting and stock price performance variables, accounting returns and stock returns, average 2.5 percent and 10.2 percent, respectively. The standard deviations of accounting returns and stock returns are 5.4 percent and 30.3 percent, respectively. Means and standard deviations are similar to those of the S&P 500 over the years 1959–1995. In addition, the (unreported) cross-sectional average firm beta, which we calculate using annual data over the 37 years, does not significantly differ from 1.

Panel B of Table 1 presents the simple correlations among our major variables. As expected, cash and total compensation are positively correlated, and both measures of compensation are positively correlated with stock price performance. Surprisingly, both measures of compensation are negatively contemporaneously correlated with accounting performance.[4] Finally, as expected, accounting performance is positively correlated with stock price performance.

Table 2 presents evidence on the time-series behavior of the compensation, accounting performance, and stock price performance data. Panel A shows that the autocorrelations for stock price performance are insignificant at all lags. In contrast, both measures of compensation and accounting performance are autocorrelated. The autocorrelations for cash compensation and accounting performance slowly damp out by lag four, although the autocorrelation for total compensation is still significant at lag four. Thus, both compensation and accounting performance have important dynamic components.

Panel B of Table 2 presents the results of unit root tests (Dickey and Fuller 1981). The null hypothesis of a unit root is rejected for all four series. This means that, over the sample period, these variables have not been subjected to changes that remain permanently in the level of the variables. These results suggest that it would be inappropriate to fit these compensation and performance data to a model that implicitly assumes unexpected performance leads to permanent changes in CEO compensation.

The 30 firms in our sample have a total of 128 different CEOs during the 1959–1995 period, resulting in an average of 4.27 different CEOs per firm, with a minimum of 2 and a maximum of 8. The average CEO tenure is 8.67 years, similar to the Hall and Liebman (1998) sample. Finally, almost all of the CEOs in our sample are promoted from within, similar to Gibbons and Murphy's (1992) sample.

IV. EMPIRICAL ANALYSIS

Long-Run Estimates from the Traditional Compensation Model

We begin by estimating a version of the traditional compensation model presented in Equation (2′). We denote CEO compensation for firm i in period t as C_{it}. Next, we introduce our measures of firm performance for firm i in period t: unexpected accounting performance, denoted as z^A_{it}, and unexpected stock price performance, denoted as z^S_{it}. We also add exogenous variables to control for broader industry-level events and trends: firm size, as measured by net sales (SIZE), and an indicator variable that takes on a value of 1 during the year of a CEO turnover, and 0 otherwise (ΔCEO). These two variables allow the level of pay to shift when the size of the firm changes or a new CEO starts on the job. Our empirical version of the traditional compensation model can then be written as:

$$C_{it} = \beta_{i0} + \rho C_{it-1} + \beta_S z^S_{it} + \beta_A z^A_{it} + \beta_{i1}\Delta CEO_{it} + \beta_{i2} SIZE_{it-1} + z^C_{it}. \qquad (3)$$

In this model, β_S, and β_A are estimated across firms, but the remaining β coefficients are

---

[4] Antle and Smith (1986) show a negative coefficient on accounting performance in their compensation regressions for a subset of their sample firms.

TABLE 2
Time-Series Behavior of CEO Compensation, Accounting Performance, and Stock Price Performance[a]
(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Autocorrelations[b]

		Lag 1	*Lag 2*	*Lag 3*	*Lag 4*
Cash Compensation	Mean:	0.703	0.549	0.409	0.347
	Std. Deviation	0.172	0.207	0.230	0.201
	Minimum	0.229	0.066	−0.079	−0.009
	Maximum	0.882	0.799	0.737	0.690
Total Compensation	Mean:	0.643	0.553	0.498	0.403
	Std. Deviation	0.201	0.208	0.124	0.163
	Minimum	−0.108	−0.247	0.253	−0.097
	Maximum	0.865	0.808	0.673	0.611
Accounting Performance	Mean:	0.696	0.446	0.288	0.235
	Std. Deviation	0.114	0.164	0.191	0.179
	Minimum	0.435	0.138	−0.173	−0.202
	Maximum	0.849	0.713	0.642	0.602
Stock Performance	Mean:	−0.048	−0.137	−0.028	0.118
	Std. Deviation	0.128	0.153	0.100	0.162
	Minimum	−0.322	−0.616	−0.201	−0.291
	Maximum	0.159	0.179	0.169	0.444

Panel B: Unit Root Tests[c]

	$\hat{\eta}$	*p-value*
Cash Compensation	0.610	0.000
Total Compensation	0.578	0.000
Accounting Performance	0.632	0.000
Stock Performance	−0.281	0.000

---

[a] We measure cash compensation as the log of the sum of all compensation paid in cash. We measure total compensation as the log of the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. We measure accounting performance as earnings before extraordinary items, scaled by average total assets and then standardized. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends. We adjust all variables for inflation.

[b] The autocorrelations in Panel A are cross-sectional averages of firm-by-firm autocorrelations.

[c] We test the null hypothesis of the unit root (i.e., $\eta = 1$) via augmented Dickey and Fuller (1981) tests (ADF). The test equation for the ADF test is $\Delta y_{it} = a_{0i} + a_1 t + (\eta - 1)y_{it-1} + a_2 \Delta y_{it-1} + a_3 \Delta y_{it-2} + \varepsilon_{it}$ where y_{it} is each individual series in turn. We perform these tests using a standard fixed-effect model (i.e., firm-specific intercepts, all other parameters equated across firms) for each series.

estimated as firm specific parameters. β_S and β_A are of particular interest because they measure the current-period response of CEO compensation to unexpected stock and accounting price performance. We estimate one version of the model with $\rho = 1$, and a second version of the model with ρ freely estimated.

z^S_{it} and z^A_{it} represent the effects of unexpected actions or events on stock price performance and accounting performance, respectively. An example of a z^S_{it} is a CEO action that increases expected future earnings, but has little or no effect on current earnings. Examples would include a CEO's unanticipated decision to increase future research and development efforts on new product lines, or to dramatically change the firm's strategic focus over the next few years. A simple example of a z_{it}^A is unobserved earnings management that inflates accounting performance in the current period. Another, more complicated example, would be an unanticipated increase in current earnings as a result of the CEO's unobserved decision to engage in a marketing plan that works better than anyone expected. In this latter case, although the action primarily affects current earnings, it may also affect expected future earnings and therefore stock prices. As we will show below, our model handles this complication by allowing unexpected accounting performance to affect stock prices contemporaneously.

Before we can estimate Equation (3), we must first obtain estimates of z_{it}^A and z_{it}^S. We model unexpected accounting performance as the residual from the following regression model:

$$A_{it} = \phi_{i0} + \phi_1 A_{it-1} + z_{it}^A, \tag{4}$$

where we denote accounting performance for firm i in period t as A_{it}. We estimate the ϕ_{i0} as firm-specific constants, and we estimate ϕ_1 across firms. The parameter ϕ_1 describes the relation between accounting performance in periods t and t−1. For example, if $\phi_1 = 1$, accounting performance follows a random walk.[5] Otherwise, if $0<\phi_1<1$, the effects of past accounting performance on current accounting performance damp out over time.

After obtaining the $\hat{z}_{it}^A$ estimated from Equation (4), we model unexpected stock price performance as the residual from the following regression:

$$S_{it} = \nu_{i0} + \nu_{i1}MKT_t + \nu_2 \hat{z}_{it}^A + z_{it}^S, \tag{5}$$

where we denote stock price performance for firm i in period t as S_{it}, and the value-weighted market return with dividends as MKT_t. We estimate the ν_{i0} as firm-specific constants, and we estimate the ν_{i1} as firm-specific parameters corresponding to firm-specific annual betas from the capital asset pricing model. The parameter ν_2 is the estimate across all firms of the effect of unexpected accounting performance on stock price performance.

After obtaining the $\hat{z}_{it}^S$ estimated from Equation (5), we use the $\hat{z}_{it}^A$ and $\hat{z}_{it}^S$ to estimate Equation (3). We denote Equations (3), (4), and (5), with ρ and ϕ restricted to be 1, as the restricted traditional model. We estimate these three equations as seemingly unrelated regressions (SURs).

As we note in Section I, prior research has concluded that: (1) increases in accounting performance positively and significantly affect current CEO compensation, and (2) accounting performance often appears to have more impact on CEO compensation than does stock price performance. Panel A of Table 3 shows that when we use the traditional model that restricts ρ and ϕ_1 to be 1, our data confirm these conclusions: (1) the coefficient on unexpected accounting performance, β_A, is positive and significant, and (2) the coefficient on unexpected accounting performance, β_A, is two to three times larger than the coefficient on unexpected stock price performance, β_S. A Wald test of the hypothesis that $\beta_A = \beta_S$ is rejected for both cash and total compensation. Thus, our data are representative in that, when we use the traditional form of the compensation model, we obtain results similar to those found in previous research.

---

[5] Although most prior compensation research uses a random walk to model the time-series behavior of annual accounting performance, Baber et al. (1998, 1999) adopt an IMA (1,1) characterization of accounting performance.

TABLE 3
Restricted Regressions of CEO Compensation on Unexpected Accounting Performance and Unexpected Stock Price Performance Based on the Traditional Model[a]
(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Results from Compensation Regressions

	Coefficients on		
	C_{t-1}	$\hat{z}_t^S$	$\hat{z}_t^A$
Cash Compensation	1.00*	0.021†	0.066†
Total Compensation	1.00*	0.033†	0.068†

Panel B: Results from Supporting Regressions

Coefficient on $\hat{z}_t^A$ in Stock Price Performance Regression	0.071†
Coefficient on A_{t-1} in Accounting Performance Regression	1.00*

† Indicates a p-value less than 0.05.

* Indicates the parameter is restricted to the reported value.

[a] We estimate the traditional compensation model in Equation (3) with $\rho = 1$:

$$C_{it} = \beta_{i0} + C_{it-1} + \beta_S \hat{z}_{it}^S + \beta_A \hat{z}_{it}^A + \beta_{i1} \Delta CEO_{it} + \beta_{i2} SIZE_{it-1} + z_{it}^C,$$

where C_{it} is the log of cash or total CEO compensation for firm i in year t. Cash compensation is the sum of all compensation paid in cash. Total compensation is the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. $\hat{z}_{it}^A$ is unexpected accounting performance for firm i in year t, estimated as the residual from a restricted version of the accounting performance regression model in Equation (4): $A_{it} = A_{it-1} + z_{it}^A$, where A_{it} is accounting performance for firm i in year t, measured as earnings before extraordinary items scaled by average total assets and then standardized to have a mean of zero and standard deviation of 1. $\hat{z}_{it}^S$ is unexpected stock price performance for firm i in year t, estimated as the residual from the stock price performance regression model in Equation (5): $S_{it} = \nu_{i0} + \nu_{i1} MKT_t + \nu_2 \hat{z}_{it}^A + z_{it}^S$, where S_{it} is stock price performance for firm i in year t, measured as the *ex post* annual rate of return to shareholders, including dividends, standardized to have a mean of zero and standard deviation of 1. We adjust all variables for inflation. We do not report the firm-specific coefficients on the constants and exogenous variables. These results are available from the authors upon request.

Figure 1 graphs the long-run response of CEO compensation to unexpected accounting and stock price performance, under the restricted traditional model in Equations (3), (4), and (5), with $\rho = 1$ and $\phi_1 = 1$. The horizontal axis measures time. The vertical axis measures the percent increase in compensation from a one-standard-deviation increase in unexpected performance (i.e., the $\partial C_{t+k}/\partial z_t^A$ and $\partial C_{t+k}/\partial z_t^S$ for $k = 0,\ldots,9$.[6] As discussed in Section II, the restricted traditional model implies that compensation responses are permanent. Thus, the current-period coefficients reported in Table 3 determine the long-run effect. For example, Panel A of Table 3 shows that a one-standard-deviation increase in unexpected accounting performance leads to a 6.8 percent increase in current-period total CEO compensation. As shown in Panel B of Figure 1, the restricted traditional model assumes that this 6.8 percent raise remains permanently in the CEO's total pay. Similarly, Panel A of Table 3 shows that a one-standard-deviation increase in unexpected stock price

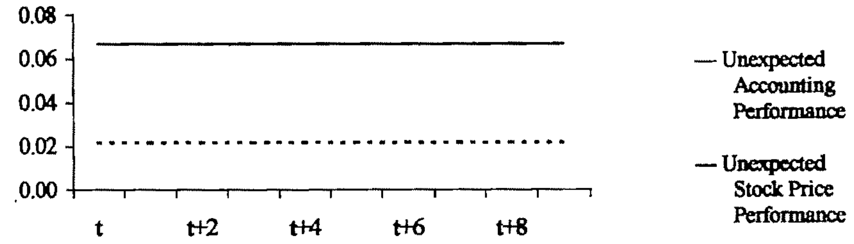
---

[6] These partial derivatives are actually taken with respect to the $\hat{z}_t^A$ and $\hat{z}_t^S$, but for notational simplicity, we suppress the hats.

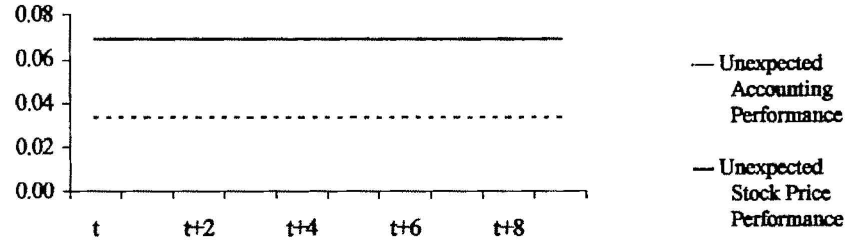
FIGURE 1

Long-Run Responses of CEO Compensation to Unexpected Accounting Performance and Unexpected Stock Price Performance from the Restricted Traditional Model[a,b,c]
(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Cash Compensation



Panel B: Total Compensation



---

[a] We measure cash compensation as the log of the sum of all compensation paid in cash. We measure total compensation as the log of the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. We measure accounting performance as earnings before extraordinary items, scaled by average total assets, and then standardized. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends, and then standardized. We adjust all variables for inflation.

[b] We estimate the restricted version of the traditional compensation model in Equation (3): $C_{it} = \beta_{i0} + C_{it-1} + \beta_S \hat{z}^S_{it} + \beta_A \hat{z}^A_{it} + \beta_{i1} \Delta CEO_{it} + \beta_{i2} SIZE_{it-1} + z^C_{it}$, where $\hat{z}^A_{it}$ is unexpected accounting performance, estimated as the residual from the restricted version of the traditional accounting performance model in Equation (4): $A_{it} = A_{it-1} + \hat{z}^A_{it}$. $\hat{z}^S_{it}$ is unexpected stock price performance, estimated as the residual from the stock price performance model in Equation (5): $S_{it} = \nu_{i0} + \nu_{i1} MKT_t + \nu_2 \hat{z}^A_{it} + \hat{z}^S_{it}$. We calculate the long-run compensation responses to unexpected stock price performance and unexpected accounting performance as the sequences of derivatives $\partial C_{t+k}/\partial z^S_t$ and $\partial C_{t+k}/\partial z^A_t$ for k = 0,...,9. The compensation responses for the traditional model shown in this table are all equal to the current-period responses, which are the coefficients on $\hat{z}^A_t$ and $\hat{z}^S_t$ in Table 3.

[c] The horizontal axis measures time (i.e., periods t, t+1, t+2, and so on). The vertical axis measures the percent increase in compensation from a one-standard-deviation increase in unexpected performance (i.e., the $\partial C_{t+k}/\partial z^S_t$ and the $\partial C_{t+k}/\partial z^A_t$).

performance leads to a 3.3 percent increase in current-period total CEO compensation. Again, the traditional model implies that this 3.3 percent raise remains permanently in the CEO's pay, as shown in Panel B of Figure 1.

In Table 4, we test how relaxing the restrictions that $\rho = 1$ in Equation (3) and $\phi_1 = 1$ in Equation (4) affects the long-run implications of the traditional model. That is, we estimate all three equations without placing any restrictions on ρ and ϕ_1. Table 4, Panel A, shows that our estimates of ρ (the coefficient on C_{t-1} in Equation (3)) are well below 1.

TABLE 4
Unrestricted Regressions of CEO Compensation on Unexpected Accounting Performance and Unexpected Stock Price Performance Based on the Traditional Model[a]
(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Relevant Results from Compensation Regressions

	Coefficients on		
	C_{t-1}	$\hat{\zeta}^S_t$	$\hat{\zeta}^A_t$
Cash Compensation	0.662†	0.022†	0.065†
Total Compensation	0.513†	0.029†	0.071†

Panel B: Relevant Results from Supporting Regressions

Coefficient on $\hat{\zeta}^A_t$ in Stock Price Performance Regression	0.060†
Coefficient on A_{t-1} in Accounting Performance Regression	0.705†

---

† Indicates a p-value less than 0.05.

[a] We estimate the unrestricted version of the traditional compensation model in Equation (3):

$$C_{it} = \beta_{i0} + \rho C_{it-1} + \beta_S \hat{\zeta}^S_{it} + \beta_A \hat{\zeta}^A_{it} + \beta_{i1} \Delta CEO_{it} + \beta_{i2} SIZE_{it-1} + \zeta^C_{it},$$

where C_{it} is the log of cash or total CEO compensation for firm i in year t. Cash compensation is the sum of all compensation paid in cash. Total compensation is the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. $\hat{\zeta}^A_{it}$ is unexpected accounting performance for firm i in year t, estimated as the residual from an unrestricted version of the accounting performance model in Equation (4): $A_{it} = \phi_{i0} + \phi_1 A_{it-1} + \hat{\zeta}^A_{it}$, where A_{it} is accounting performance for firm i in year t, measured as earnings before extraordinary items scaled by average total assets and then standardized to have a mean of zero and standard deviation of 1. $\hat{\zeta}^S_{it}$ is unexpected stock price performance for firm i in year t, estimated as the residual from the stock price performance model in Equation (5): $S_{it} = \nu_{i0} + \nu_{i1} MKT_t + \nu_2 \hat{\zeta}^A_{it} + \hat{\zeta}^S_{it}$, where S_{it} is stock price performance for firm i in year t, measured as the *ex post* annual rate of return to shareholders, including dividends, standardized to have a mean of zero and standard deviation of 1. We adjust all variables for inflation. We do not report the firm-specific coefficients on the constants and exogenous variables. These results are available from the authors upon request.

For total compensation the estimated value of ρ is only 0.513. A Wald test of the hypothesis that $\rho = 1$ is rejected in both the cash and total compensation models. Panel B shows that the estimate of ϕ_1 (the coefficient on A_{t-1} in Equation (4)) is 0.705, and the hypothesis that $\phi_1 = 1$ is rejected. Overall, the traditional model restrictions of $\rho = 1$ and $\phi_1 = 1$ are *not* consistent with the data. This finding implies that models of the pay-performance relation that use first-differenced compensation and first-differenced accounting performance are misspecified.

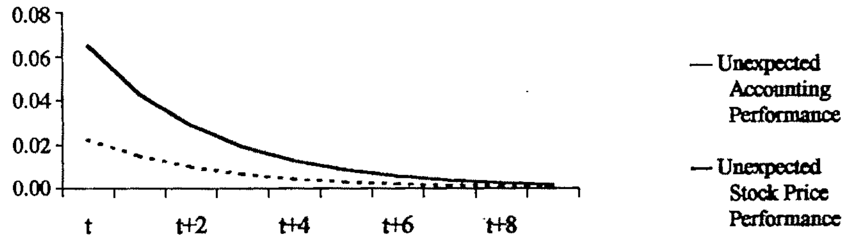
In Figure 2 we show that relaxing the restrictions that $\rho = 1$ and $\phi_1 = 1$ has substantial implications for the long-run response of CEO compensation to unexpected firm performance. Because Table 4 shows that $\rho < 1$ and $\phi_1 < 1$, the CEO's raise for unexpectedly good firm performance no longer remains in his compensation permanently. Instead, the CEO's raise dissipates over time. For example, although Panel A of Table 4 shows that the CEO receives an initial raise in total compensation of 7.1 percent for an unexpected increase in accounting performance in the current period, only about 1 percent (i.e., $\rho^3 \times \beta_A = (0.513)^3 \times 0.071$) remains in his pay by year t+3.

As explained above, the model in Equation (3) uses the single parameter, ρ, to describe the dynamics of the compensation responses. For values of ρ greater than zero but less than

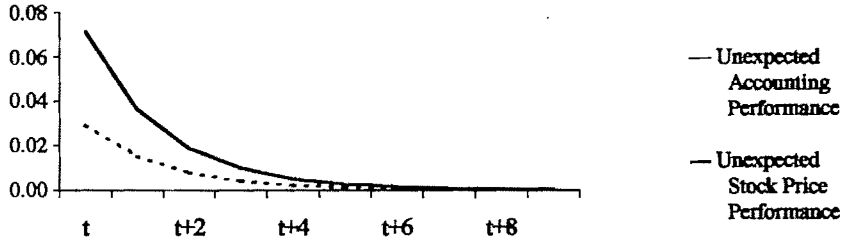
FIGURE 2

Long-Run Responses of CEO Compensation to Unexpected Accounting Performance and Unexpected Stock Price Performance from the Unrestricted Traditional Model[a,b,c]
(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Cash Compensation



Panel B: Total Compensation



---

[a] We measure cash compensation as the log of the sum of all compensation paid in cash. We measure total compensation as the log of the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. We measure accounting performance as earnings before extraordinary items, scaled by average total assets, and then standardized. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends, and then standardized. We adjust all variables for inflation.

[b] We estimate the unrestricted version of the traditional compensation model in Equation (3): $C_{it} = \beta_{i0} + \rho C_{it-1} + \beta_S \hat{z}_{it}^S + \beta_A \hat{z}_{it}^A + \beta_{i1} \Delta CEO_{it} + \beta_{i2} SIZE_{it-1} + z_{it}^C$, where $\hat{z}_{it}^A$ is unexpected accounting performance, estimated as the residual from the unrestricted version of the traditional accounting performance model in Equation (4): $A_{it} = \phi_{i0} + \phi_1 A_{it-1} + \hat{z}_{it}^A$. $\hat{z}_{it}^S$ is unexpected stock price performance, estimated as the residual from the stock price model in Equation (5): $S_{it} = \nu_{i0} + \nu_{i1} MKT_t + \nu_2 \hat{z}_{it}^A + \hat{z}_{it}^S$. We calculate the long-run compensation responses to unexpected stock price performance and unexpected accounting performance as the sequences of derivatives $\partial C_{t+k}/\partial z_t^S$ and $\partial C_{t+k}/\partial z_t^A$ for $k = 0,...,9$. The current-period compensation responses for the traditional model are the coefficients on $\hat{z}_t^A$ and $\hat{z}_t^S$ in Table 4. Over time these responses decay at a rate of ρ^k, where ρ is the coefficient on C_{it-1} in Equation (3).

[c] The horizontal axis measures time (i.e., periods t, t+1, t+2, and so on). The vertical axis measures the percent increase in compensation from a one-standard-deviation increase in unexpected performance (i.e., the $\partial C_{t+k}/\partial z_t^S$ and the $\partial C_{t+k}/\partial z_t^A$).

1, the effect of unexpected performance on CEO compensation decays over time, at a constant rate of $1 - \rho$. However, a model that assumes a constant decay rate will not uncover a more complicated pattern of responses, such as a cyclical response or the hump-shaped response found in Boschen and Smith (1995). In the next section we use a more flexible VAR approach to fully examine the current and long-run effects of performance on pay.

Long-Run Estimates from the Vector Autoregression Model

We present a vector autoregression (VAR) model of CEO compensation and firm performance that extends Boschen and Smith's (1995) model by adding accounting performance, in addition to stock price performance. This model allows for more complicated dynamics in three ways. First, we include three lags of CEO compensation, accounting performance, and stock price performance. Second, (unreported) Granger (1969) causality tests show that stock price performance and accounting performance Granger-cause one another, so we allow accounting and stock price performance to be jointly determined. Third, we allow for feedback from compensation to performance (Abowd 1990; Anderson et al. 2000; Hayes and Schaefer 2000).

The VAR model we present below contains the same set of three variables, C_{it}, A_{it}, and S_{it}, as do Equations (3), (4), and (5).[7] We continue to include the exogenous variables $SIZE_{t-1}$, ΔCEO_t, and MKT_t. In addition to these variables, we enter the lagged market return (MKT_{t-1}) in the compensation equation and the levels of current and lagged real Gross Domestic Product (GDP_t and GDP_{t-1}) in the equation for A_{it}.[8,9] The model, which is similar to Boschen and Smith's (1995) model (1) extended to include accounting performance, is:

$$\begin{aligned}
C_{it} &= \gamma_{i10} + \sum_{\ell=1}^{3} \gamma_{i11}^{\ell} C_{it-\ell} + \sum_{\ell=0}^{3} \gamma_{i12}^{\ell} S_{it-\ell} + \sum_{\ell=0}^{3} \gamma_{i13}^{\ell} A_{it-\ell} + \gamma_{iM}^{1} MKT_{t-1} \\
&\quad + \gamma_{i\Delta}^{0} \Delta CEO_{it} + \gamma_{iZ}^{1} SIZE_{it-1} + z_{it}^{C} \\
S_{it} &= \gamma_{i20} + \sum_{\ell=0}^{3} \gamma_{i21}^{\ell} C_{it-\ell} + \sum_{\ell=1}^{3} \gamma_{i22}^{\ell} S_{it-\ell} + \sum_{\ell=0}^{3} \gamma_{i23}^{\ell} A_{it-\ell} + \gamma_{iM}^{0} MKT_{t} + z_{it}^{S} \qquad (6) \\
A_{it} &= \gamma_{i30} + \sum_{\ell=0}^{3} \gamma_{i31}^{\ell} C_{it-\ell} + \sum_{\ell=0}^{3} \gamma_{i32}^{\ell} S_{it-\ell} + \sum_{\ell=1}^{3} \gamma_{i33}^{\ell} A_{it-\ell} + \gamma_{iGDP}^{0} GDP_{t} \\
&\quad + \gamma_{iGDP}^{1} GDP_{t-1} + z_{it}^{A}.
\end{aligned}$$

We continue to interpret the z_{it}^{A} and z_{it}^{S} as unexpected accounting performance and unexpected stock price performance, respectively. As before, the partial derivatives from the system in (6) (i.e., $\partial C_{t+k}/\partial z_t^A$ and $\partial C_{t+k}/\partial z_t^S$, for $k = 0, 1, 2, \ldots$) are the responses of compensation in period t+k to unexpected performance in t. The model looks different because z_{it}^{A} and z_{it}^{S} are not substituted directly into the compensation equation, as we did in Equation (3). Instead, they enter the equation for C_{it} through A_{it} and S_{it}. Thus, in the VAR we calculate the compensation responses (i.e., the partial derivatives) using a two-step process, as Appendix B explains in more detail. First, we estimate the VAR equations for the system in (6) and obtain the regression coefficients. Then we invert the estimated model to obtain its moving-average form (i.e., C_{it}, S_{it}, and A_{it} written as a function of the current and past values of z_{it}^{A}, z_{it}^{S}, z_{it}^{C}, and the exogenous variables) and calculate the partial derivatives from the moving-average form of the model.

---

[7] We modify the standard VAR by entering exogenous variables in the system. In addition, most VARs do not estimate the parameters on the current values of the endogenous variables.

[8] Adut et al. (2003) include a time-trend variable in a model of CEO compensation. They find that a deterministic trend alters the relation between CEO compensation and restructuring charges. However, including trend variables (time and time-squared) in our compensation equation does not affect our inferences.

[9] GDP plays a dual role. First, this variable controls for variations in real spending that would likely affect current and near-future accounting performance. Second, GDP, as an exogenous variable, is an additional instrument for use in the Two-Stage Least Squares (2SLS) approach we use to deal with the simultaneity in the VAR model.

We use the 2SLS approach (Hamilton 1994, 238–241) to obtain consistent estimates of the regression coefficients in (6). In this approach, we use the lagged values of compensation, stock price performance, accounting performance, and the exogenous variables as instruments to obtain fitted values for the endogenous variables.[10] Then, we substitute these fitted values for the endogenous variables (i.e., compensation, stock price performance, and accounting performance) and estimate each equation as an SUR. The confidence intervals for the compensation responses cannot be easily obtained analytically because the long-run responses are nonlinear combinations of the regression coefficients in Model (6). Thus, we obtain confidence intervals for the long-run responses using Monte Carlo simulations in which we calculate empirical distributions of the response functions based on 1,000 draws from the estimated distributions of the underlying VAR parameters (Hamilton 1994, 337).

Table 5 presents the estimation results for the model in (6). Panel A shows the results for cash compensation and Panel B shows the results for total compensation. The compensation regressions show that nearly all lags of compensation, all lags of stock price performance, and nearly all lags of accounting performance are significantly related to current compensation. The stock price performance regressions show that all lags of stock price performance, at least one lag of accounting performance, and most lags of compensation are significantly related to current stock price performance. Finally, the accounting performance regressions show that all lags of accounting performance, all lags of stock price performance, and all lags of compensation are significantly related to current accounting performance. Overall, the significance of these lags demonstrates that there are, in fact, substantial and important dynamics associated with each variable.

We use the coefficients in Table 5 to calculate the long-run responses of compensation to unexpected accounting and stock price performance that are presented in Figure 3. Specifically, we invert the system and then calculate the partial derivatives $\partial C_{t+k}/\partial z_t^A$ and $\partial C_{t+k}/\partial z_t^S$, for $k = 0,...,9$. Panel A of Figure 3 shows the compensation responses for cash compensation and Panel B of Figure 3 shows the responses for total compensation. The compensation responses to unexpected stock price performance exhibit a "hump-shaped" pattern, consistent with Boschen and Smith's (1995) results. The hump-shaped pattern means that the partial derivatives are increasing in value for the first several years, but decline toward zero in later years. In contrast, the response of CEO compensation to accounting performance is positive and statistically significant in the first period (as shown in Table 5 and in Figure 3), but the response falls off very quickly. In fact, by the year t+3, the compensation response is negative (although not statistically significant). That is, the CEO seems to reap rewards for good accounting performance for one or two periods, but then these rewards are taken back in years t+3 through t+9.

Table 6 summarizes the current-period and long-run responses of CEO compensation to unexpected accounting and stock price performance. We use this summary of the compensation responses to illustrate the rewards accruing to a hypothetical CEO in our sample who earns the sample average of $1 million in total compensation. Panel A summarizes the compensation responses indicated by the traditional restricted model, as shown in Panel A of Table 3 and in Figure 1. This traditional restricted model suggests that our hypothetical CEO would receive a reward of $68,000 in the current year's total compensation and a

---

[10] The average R^2 of the regressions that yield these fitted values are: 0.71 for the cash compensation equation, 0.69 for the total compensation equation, 0.35 for the stock price performance equation, and 0.73 for the accounting performance equation.

TABLE 5
Vector Autoregression Model of CEO Compensation, Accounting Performance, and Stock Price Performance[a,b]
(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Cash Compensation

	Dependent Variable		
Coefficients on	*Compensation*	*Stock Price Performance*	*Accounting Performance*
C_t	—	$0.464^{\dagger}$	$0.432^{\dagger}$
C_{t-1}	$0.570^{\dagger}$	$-0.226^{\dagger}$	$-0.120^{\dagger}$
C_{t-2}	$0.156^{\dagger}$	-0.048	$0.081^{\dagger}$
C_{t-3}	-0.049	$-0.173^{\dagger}$	$0.097^{\dagger}$
SR_t	$0.013^{\dagger}$	—	$0.197^{\dagger}$
SR_{t-1}	$0.026^{\dagger}$	$-0.063^{\dagger}$	$0.123^{\dagger}$
SR_{t-2}	$0.018^{\dagger}$	$-0.040^{\dagger}$	$0.073^{\dagger}$
SR_{t-3}	$0.008^{\dagger}$	$-0.045^{\dagger}$	$0.029^{\dagger}$
AR_t	$0.077^{\dagger}$	$0.210^{\dagger}$	—
AR_{t-1}	$-0.062^{\dagger}$	$-0.233^{\dagger}$	$0.525^{\dagger}$
AR_{t-2}	$0.019^{\dagger}$	-0.000	$-0.114^{\dagger}$
AR_{t-3}	$-0.027^{\dagger}$	-0.016	$0.122^{\dagger}$

Panel B: Total Compensation

	Dependent Variable		
Coefficients on	*Compensation*	*Stock Price Performance*	*Accounting Performance*
C_t	—	$0.395^{\dagger}$	$0.264^{\dagger}$
C_{t-1}	$0.389^{\dagger}$	$-0.140^{\dagger}$	$0.099^{\dagger}$
C_{t-2}	$0.210^{\dagger}$	$-0.086^{\dagger}$	$0.139^{\dagger}$
C_{t-3}	$0.212^{\dagger}$	$-0.161^{\dagger}$	$0.084^{\dagger}$
SR_t	$0.082^{\dagger}$	—	$0.175^{\dagger}$
SR_{t-1}	$0.040^{\dagger}$	$-0.067^{\dagger}$	$0.129^{\dagger}$
SR_{t-2}	$0.055^{\dagger}$	$-0.048^{\dagger}$	$0.065^{\dagger}$
SR_{t-3}	$0.018^{\dagger}$	$-0.059^{\dagger}$	$0.022^{\dagger}$
AR_t	$0.038^{\dagger}$	$0.208^{\dagger}$	—
AR_{t-1}	$-0.029^{\dagger}$	$-0.219^{\dagger}$	$0.473^{\dagger}$
AR_{t-2}	0.005	0.004	$-0.086^{\dagger}$
AR_{t-3}	$-0.024^{\dagger}$	$-0.031^{\dagger}$	$0.091^{\dagger}$

[†] Indicates a p-value less than 0.05.

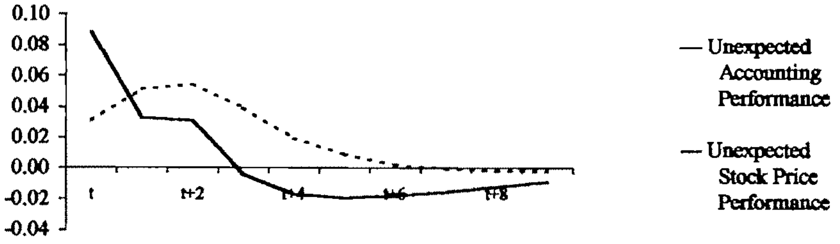
[a] We measure cash compensation as the log of the sum of all compensation paid in cash. We measure total compensation as the log of the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. We measure accounting performance as earnings before extraordinary items, scaled by average total assets, and then standardized. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends, and then standardized. We adjust all variables for inflation.

[b] We estimate the model from a set of pooled time-series seemingly unrelated regressions based on the system in (6). We do not report the coefficients on the constants and exogenous variables. These results are available from the authors upon request.

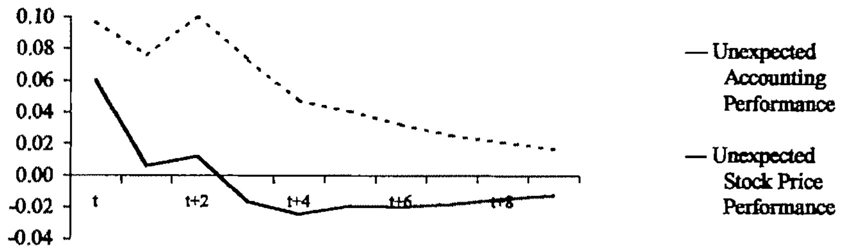
FIGURE 3

Long-Run Responses of CEO Compensation to Unexpected Accounting Performance and Unexpected Stock Price Performance Calculated from Vector Autoregression Coefficients[a,b,c] (1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Cash Compensation



Panel B: Total Compensation



---

[a] We measure cash compensation as the log of the sum of all compensation paid in cash. We measure total compensation as the log of the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. We measure accounting performance as earnings before extraordinary items, scaled by average total assets, and then standardized. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends, and then standardized. We adjust all variables for inflation.

[b] We obtain the responses of cash and total CEO compensation to unexpected accounting performance and unexpected stock price performance via a two-step process. First, we estimate a vector autoregression based on the system in (6). In the second step, we invert the system and then calculate the long-run compensation responses to unexpected stock price performance and unexpected accounting performance as the sequences of derivatives $\partial C_{t+k}/\partial z_t^S$ and $\partial C_{t+k}/\partial z_t^A$ for k = 0,...,9.

[c] The horizontal axis measures time (i.e., periods t, t+1, t+2, and so on). The vertical axis measures the percent increase in compensation from a one-standard-deviation increase in unexpected performance (i.e., the $\partial C_{t+k}/\partial z_t^S$ and $\partial C_{t+k}/\partial z_t^A$).

total of $680,000 over ten years for a one-standard-deviation increase in accounting performance. However, the traditional restricted model suggests that the CEO's total compensation reward for a one-standard-deviation increase in stock price performance is only $33,000 in the current period and $330,000 over ten years.

Panel B of Table 6 summarizes the compensation responses calculated from the VAR model (i.e., by estimating the coefficients for the model in (6) as shown in Table 5, inverting the system, and calculating the partial derivatives shown in Figure 3). Our hypothetical CEO receives an increase in total compensation of $60,000 in the current year for a one-standard deviation increase in accounting performance, but the VAR model suggests that this reward is followed by future pay cuts. The net loss to our CEO over the ten-year period

TABLE 6
Current-Period and Long-Run Cumulative Responses of CEO Compensation to Unexpected Accounting Performance and Unexpected Stock Price Performance[a]
(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Traditional Compensation Model[b]

Cash Compensation

	Current-Period[d]	*Long-Run Cumulative*[e]
Response to z_t^A	$0.066^{\dagger}$	$0.660^{\dagger}$
Response to z_t^S	$0.021^{\dagger}$	$0.210^{\dagger}$

Total Compensation

	Current-Period[d]	*Long-Run Cumulative*[e]
Response to z_t^A	$0.068^{\dagger}$	$0.680^{\dagger}$
Response to z_t^S	$0.033^{\dagger}$	$0.333^{\dagger}$

Panel B: Vector Autoregression Model[c]

Cash Compensation

	Current-Period[d]	*Long-Run Cumulative*[e]
Response to z_t^A	$0.088^{\dagger}$	0.054
Response to z_t^S	$0.031^{\dagger}$	$0.199^{\dagger}$

Total Compensation

	Current-Period[d]	*Long-Run Cumulative*[e]
Response to z_t^A	$0.060^{\dagger}$	−0.049
Response to z_t^S	$0.097^{\dagger}$	$0.529^{\dagger}$

[†] Indicates a p-value less than 0.05.

[a] We measure cash compensation as the log of the sum of all compensation paid in cash. We measure total compensation as the log of the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. We adjust all variables for inflation. We measure accounting performance as earnings before extraordinary items, scaled by average total assets, and then standardized. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends, and then standardized. We adjust all variables for inflation.

[b] In Panel A, we obtain the responses of cash and total CEO compensation to unexpected accounting performance and unexpected stock price performance by estimating the restricted traditional compensation model in Equation (3), using the estimates calculated from Equations (4) and (5). We calculate the long-run compensation responses to unexpected stock price performance and unexpected accounting performance as the sequences of derivatives $\partial C_{t+k}/\partial z_t^S$ and $\partial C_{t+k}/\partial z_t^A$ for k = 0,...,9.

[c] In Panel B, we obtain the responses of cash and total CEO compensation to unexpected accounting performance and unexpected stock price performance via a two-step process. First, we estimate a vector autoregression based on the system in (6). In the second step, we invert the system and calculate the long-run compensation responses to unexpected stock price performance and unexpected accounting performance as the sequences of derivatives $\partial C_{t+k}/\partial z_t^S$ and $\partial C_{t+k}/\partial z_t^A$ for k = 0,...,9.

[d] For both models, we calculate the current-period responses of cash and total CEO compensation to unexpected stock price performance and unexpected accounting performance as $\partial C_t/\partial z_t^S$ and $\partial C_t/\partial z_t^A$.

[e] For both models, we calculate the cumulative responses of cash and total CEO compensation to unexpected stock price performance and unexpected accounting performance as the sums of the derivatives above, that is, $\Sigma_{k=0}^{9}\ \partial C_{t+k}/\partial z_t^S$ and $\Sigma_{k=0}^{9}\ \partial C_{t+k}/\partial z_t^A$, respectively.

is $49,000 (a value not significantly different from zero). In contrast, the VAR model suggests that our CEO's reward for a one-standard-deviation increase in stock price performance is $97,000 in the current period and $529,000 over ten years.

We now consider whether this puzzling dynamic effect of unexpected accounting performance on CEO compensation can be explained by the three major hypotheses about why firms use accounting performance data in setting CEO compensation. As discussed in Section II, the incentive, filtering, and effort allocation hypotheses can be consistent with either positive or negative relations between CEO compensation and accounting performance. However, it is difficult to see how using accounting performance in compensation contracts to provide incentives and/or to filter noise could explain a response of compensation to accounting performance that *shifts* from positive to negative. Use of accounting performance as a straightforward effort incentive cannot explain this pattern because the dynamics imply that rewards for actions that raise current accounting performance will be followed by (future) penalties for these same actions. These dynamics also appear inconsistent with the use of accounting performance to filter noise, because such a pattern would require that the correlations between current unexpected accounting performance and future realizations of unexpected stock price performance reverse sign over time.

One possible explanation is that incentive effects dominate in the short run (consistent with the positive compensation responses to accounting performance) and filtering effects dominate in the long run (consistent with the negative responses). However, to play a filtering role over the long run, accounting data lagged two or more years must convey substantial information about the current noise in stock prices. Although this is possible, accounting data are not usually considered strongly predictive of variation in future stock price performance.

Turning to the effort allocation hypothesis, one possible explanation is that firms use accounting performance to direct CEOs *away* from certain types of actions (Bushman and Indjejikian 1993; Lambert 2001). For example, consider a firm that ratchets performance standards (Indjejikian and Nanda 1999; Murphy 2000). In such a firm, a CEO who takes an action resulting in an earnings increase that is not sustainable is rewarded in the short run, but penalized in the long run. As another example, consider a CEO who chooses to manage earnings upward by aggressively accelerating the recognition of revenues or deferring the recognition of expenses. Such a choice may have no cash flow consequences overall, but the acceleration of earnings into the current period will, by definition, result in future earnings that are lower than they would otherwise be. Thus, a positive initial compensation response is followed by a later adverse response. Both of these examples suggest compensation responses to unexpected accounting performance that are initially positive, but decrease and become negative over time, consistent with our findings.

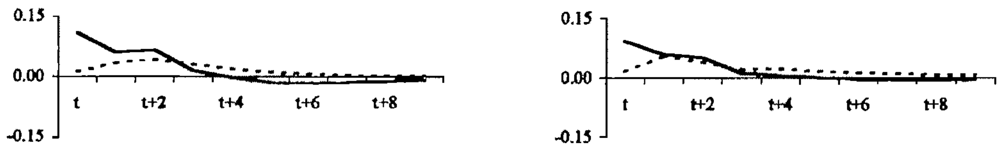
Robustness

We conduct a variety of robustness tests on the VAR model in (6) to evaluate whether a more parsimonious model can adequately capture the full dynamic responses of CEO compensation to firm performance. For brevity, we summarize our findings but do not report tabulated results. We begin by testing the sensitivity of the compensation responses shown in Table 5 to simplifications of the performance equations in (6). First, we eliminate the feedback from compensation back to both performance equations (i.e., we set $\gamma_{21}^{\ell} = \gamma_{31}^{\ell} = 0$ for all ℓ). Second, we eliminate the feedback from stock price performance in the accounting performance equation (i.e., we set $\gamma_{32}^{\ell} = 0$ for all ℓ). These restrictions reduce the model to a recursive form, which eliminates the need for 2SLS estimation. Third, we limit the performance equations to a maximum of one lag of any variable (i.e., $\gamma_{22}^{2} = \gamma_{22}^{3} = \gamma_{23}^{2} = \gamma_{23}^{3} = \gamma_{33}^{2} = \gamma_{33}^{3} = 0$). In each case, despite the increasing restrictions placed on these equations, the shape and the magnitude of the compensation responses remain virtually identical to the main VAR results in Figure 3. Thus, we can model firm performance with fewer parameters and still capture the essential compensation dynamics.

In contrast, when we restrict the compensation equation, we find substantial changes in the magnitude of the compensation responses. Specifically, when we eliminate all but

FIGURE 4
Long-Run Responses of CEO Compensation to Unexpected Accounting Performance and Unexpected Stock Price Performance across Three Decades[a,b]
(1,110 firm-years—30 firms for the 1959–1995 period)

Panel A: Compensation Responses for the Period 1960–1971[e]



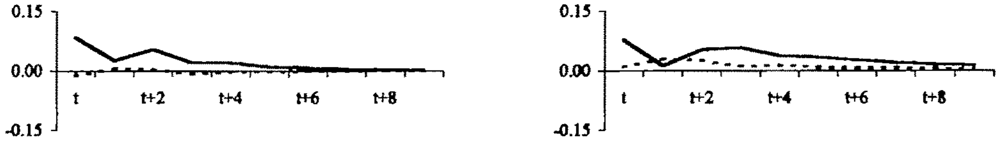
Cash Compensation

	Current-Period[c]	*Long-Run Cumulative*[d]
Response to z_t^A	$0.108^{\dagger}$	$0.181^{\dagger}$
Response to z_t^S	$0.013^{\dagger}$	$0.167^{\dagger}$

Total Compensation

	Current-Period[c]	*Long-Run Cumulative*[d]
Response to z_t^A	$0.092^{\dagger}$	$0.194^{\dagger}$
Response to z_t^S	$0.015^{\dagger}$	$0.205^{\dagger}$

Panel B: Compensation Responses for the Period 1972–1983[e]



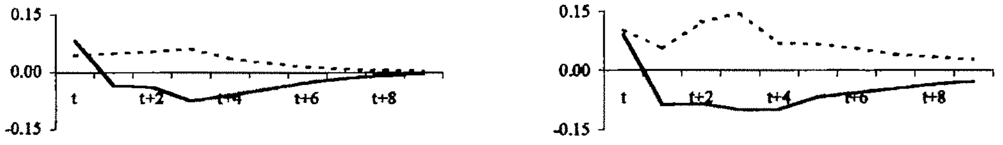
Cash Compensation

	Current-Period[c]	*Long-Run Cumulative*[d]
Response to z_t^A	$0.084^{\dagger}$	$0.227^{\dagger}$
Response to z_t^S	-0.012	$-0.022^{\dagger}$

Total Compensation

	Current-Period[c]	*Long-Run Cumulative*[d]
Response to z_t^A	$0.078^{\dagger}$	$0.341^{\dagger}$
Response to z_t^S	0.010	$0.120^{\dagger}$

Panel C: Compensation Responses for the Period 1984–1995[e]



Cash Compensation

	Current-Period[c]	*Long-Run Cumulative*[d]
Response to z_t^A	$0.082^{\dagger}$	$-0.229^{\dagger}$
Response to z_t^S	$0.043^{\dagger}$	$0.295^{\dagger}$

Total Compensation

	Current-Period[c]	*Long-Run Cumulative*[d]
Response to z_t^A	$0.092^{\dagger}$	$-0.510^{\dagger}$
Response to z_t^S	$0.104^{\dagger}$	$0.720^{\dagger}$

(*continued on next page*)

FIGURE 4 (continued)

---

† Indicates a p-value less than 0.05.

[a] We measure cash compensation as the log of the sum of all compensation paid in cash. We measure total compensation as the log of the sum of cash compensation, stock grants, stock option grants, and any other noncash compensation. We measure accounting performance as earnings before extraordinary items, scaled by average total assets, and then standardized. We measure stock price performance as the *ex post* annual rate of return to shareholders, including dividends, and then standardized. We adjust all variables for inflation.

[b] We obtain the responses of cash and total CEO compensation to unexpected accounting performance and unexpected stock price performance via a two-step-process. First, we estimate a vector autoregression based on a recursive version of the system in (6) with the performance equations restricted to only one lag. We added time-period dummy interactions for the constants and for each performance variable (current and lagged). We use two dummies, one to indicate the 1972–1983 period, and the other to indicate the 1984–1995 period. In the second step, we invert the system and calculate the long-run compensation responses to unexpected stock price performance and unexpected accounting performance as the sequences of derivatives $\partial C_{t+k}/\partial z_t^S$ and $\partial C_{t+k}/\partial z_t^A$ for $k = 0,\ldots,9$.

[c] We calculate the current-period responses of cash and total CEO compensation to unexpected stock price performance and unexpected accounting performance as $\partial C_t/\partial z_t^S$ and $\partial C_t/\partial z_t^A$.

[d] We calculate the cumulative responses of cash and total CEO compensation to unexpected stock price performance and unexpected accounting performance as the sums of the derivatives above, that is, $\Sigma_{k=0}^{9}\ \partial C_{t+k}/\partial z_t^S$ and $\Sigma_{k=0}^{9}\ \partial C_{t+k}/\partial z_t^A$, respectively.

[e] The horizontal axis of each graph measures time (i.e., periods t, t+1, t+2, and so on). The vertical axis measures the percent increase in compensation from a one-standard-deviation increase in unexpected performance (i.e., the $\partial C_{t+k}/\partial z_t^S$ and the $\partial C_{t+k}/\partial z_t^A$).

one lag of compensation in the compensation equation, the cumulative response to unexpected stock price performance drops by about half. When we further restrict the compensation equation to include only one lag of all the variables, much of the long-run compensation response is eliminated altogether. These results suggest that it is necessary to include multiple lags of compensation, accounting performance, and stock price performance in the compensation equation, in order to fully capture the dynamic compensation responses.

Changing Dynamics over Three Decades

Figure 4 shows how the responses of compensation to unexpected accounting performance and unexpected stock price performance have changed over time. To obtain these estimates, we use a version of the system in (6) that includes time-period dummy variables interacted with the firm-specific constants and interacted with each performance variable. We use two dummies, one to indicate the 1972–1983 period, and the other to indicate the 1984–1995 period. Because of the large number of parameters involved, we take advantage of the findings from our robustness testing above. Specifically, we simplify the performance equations in Equation (6) so that (1) stock price performance is only affected by one lag of stock price performance and current and one lag of accounting performance, and (2) accounting performance is only affected by one lag of accounting performance. Panels A, B, and C of Figure 4 show the responses of CEO compensation to unexpected accounting performance and unexpected stock price performance for the 1960-1971, 1972–1983, and 1984–1995 periods, respectively.

Panels A, B, and C of Figure 4 show that the current-period effect of unexpected accounting performance is stable over the three decades. In contrast, the current-period effect of unexpected stock price performance rises considerably in the 1984–1995 period. These latter findings are consistent with those reported in Murphy (1999) and Bushman et al. (2000). The long-run effects of unexpected stock price performance have also increased over time. However, the long-run effects of unexpected accounting performance have declined sharply in the last decade and, in fact, have shifted from positive to significantly negative. These results are consistent with the idea that firms have implicitly shifted their CEO compensation arrangements to reward stock price performance and de-emphasize accounting performance. Future research on how firms use accounting performance to provide incentives for specific types of actions seems a promising avenue in explaining these shifts.

V. CONCLUSION

Prior empirical research using restricted models of the pay-performance relation finds that CEO compensation often responds more strongly to accounting performance than to stock price performance (Antle and Smith 1986; Lambert and Larcker 1987; Rosen 1992; Sloan 1993). However, we show that this finding does not hold when the empirical model is extended to allow unrestricted estimation of the long-run response of CEO compensation to unexpected accounting performance. We find that the compensation response to unexpected stock price performance is positive and persistent over time, but the compensation response to unexpected accounting performance is positive in the early years, and negative in later years. Overall, the cumulative long-run response of CEO compensation to unexpected accounting performance is not significantly different from zero.

From a methodological viewpoint, we conclude that pay-performance models should include at least three lags of all compensation and performance variables in order to fully capture the dynamics of the pay-performance relation. The traditional compensation model based on first-differences of compensation and accounting performance is, therefore, misspecified. A limitation of our study is that it is based on only 30 firms (even though there are 1,110 firm-years of data). Also, requiring that firms have 37 years of data introduces survivorship bias, so the dynamic pay-performance relation we describe generalizes only to large, mature, surviving firms.

Finally, it is tempting to interpret our findings as diminishing the empirical importance of accounting performance in executive compensation arrangements. However, we have argued above that CEOs who take actions that increase accounting performance only temporarily, may face future penalties. These penalties may arise because firms ratchet performance standards, or because earnings management that accelerates earnings into the current period necessarily trades off current against future earnings. We conclude that understanding the long-run relation between accounting performance and compensation will require a complex set of theoretical considerations and more empirical work.

APPENDIX A
Sample Firms and Industry

Firm	*Industry*
1. Allied Signal	Manufacturing (Diversified)
2. Boeing	Aerospace/Defense
3. Corning	Manufacturing (Diversified)
4. Dow Chemical	Chemicals
5. Eastman Kodak	Photography/Imaging
6. Eaton	Manufacturing (Diversified)
7. Federal Mogul	Auto Parts and Equipment

8. FMC	Chemicals (Diversified)
9. Ford Motor	Automobiles
10. Goodrich (B. F.)	Chemicals (Diversified)
11. Grace (W. R.)	Chemicals (Specialty)
12. Harsco	Manufacturing (Diversified)
13. Hercules	Foods
14. Hershey	Foods
15. Kellogg	Foods
16. Kerr McGee	Oil and Gas (Exploration/Production)
17. Kimberly Clark	Household Products (Nondurable)
18. Minnesota Mining	Manufacturing (Diversified)
19. Mobil	Oil (Internationally Integrated)
20. Monsanto	Chemicals (Diversified)
21. NL Industries	Industrial Inorganic Chemicals
22. Northrop Grumman	Aerospace/Defense
23. Pfizer	Health Care (Drugs/Pharmaceuticals)
24. PPG Industries	Chemicals (Diversified)
25. Raytheon	Electronics/Defense
26. Rockwell International	Electrical Equipment
27. Rohm & Haas	Chemicals
28. Tenneco	Manufacturing (Diversified)
29. TRW	Auto Parts and Equipment
30. Union Carbide	Chemicals

APPENDIX B

Calculation of Dynamic Responses for Vector Autoregression Model of CEO Compensation, Accounting Performance, and Stock Price Performance

In this appendix, we detail the calculation of the dynamic responses for the vector autoregression model in (6). We can write this model, with constants, exogenous variables, and firm indicators suppressed, as:

$$\begin{bmatrix} 1 & -\gamma_{12}^{(0)} & -\gamma_{13}^{(0)} \\ -\gamma_{21}^{(0)} & 1 & -\gamma_{23}^{(0)} \\ -\gamma_{31}^{(0)} & -\gamma_{32}^{(0)} & 1 \end{bmatrix} \begin{bmatrix} C_t \\ S_t \\ A_t \end{bmatrix} = \sum_{\ell=1}^{3} \begin{bmatrix} \gamma_{11}^{(\ell)} & \gamma_{12}^{(\ell)} & \gamma_{13}^{(\ell)} \\ \gamma_{21}^{(\ell)} & \gamma_{22}^{(\ell)} & \gamma_{23}^{(\ell)} \\ \gamma_{31}^{(\ell)} & \gamma_{32}^{(\ell)} & \gamma_{33}^{(\ell)} \end{bmatrix} \begin{bmatrix} C_{t-\ell} \\ S_{t-\ell} \\ A_{t-\ell} \end{bmatrix} + \begin{bmatrix} z_t^C \\ z_t^S \\ z_t^A \end{bmatrix} \tag{B1}$$

and condense it into vector form as:

$$\Gamma_0 Y_t = \sum_{\ell=1}^{3} \Gamma_\ell Y_{t-\ell} + Z_t, \tag{B2}$$

where $Y_t' = [C_t \quad S_t \quad A_t]$, $Z_t' = [z_t^C \quad z_t^S \quad z_t^A]$, and the Γ_ℓ are the coefficient matrices on the endogenous vectors at lag ℓ.

One standard version of the vector autoregression assumes that $\Gamma_0 = I$ and imposes a matrix of restrictions on the Z_t (which imposes an ordering on the estimation of the vector autoregression). Unlike the standard vector autoregression, the model we estimate does not restrict Γ_0 to identity. Instead, we allow for simultaneity to accommodate the endogenous relations among compensation, stock price performance, and accounting performance. Our model also differs from the standard vector autoregression by recognizing that the three dependent variables do not exist as a fully closed system. Thus, our model allows exogenous variables to affect the system. However, since the exogenous variables do not affect the calculation of the dynamic responses, we suppress them in this appendix.

To calculate the dynamic responses, we invert Equation (B2) to obtain the moving-average form shown in Equation (B3):

$$Y_t = \sum_{k=0}^{t} \beta_k Z_{t-k}, \tag{B3}$$

where the β_k are complex functions of the Γ_t matrices. The dynamic response of compensation through period t+k to unexpected stock price performance in period t is the sequence of derivatives $\partial C_{t+k}/\partial z_t^S$ that correspond to the elements $\beta_k(1,2)$, for $k = 0,\ldots,K$. Analogously, the dynamic response of compensation through period t+k to unexpected accounting performance in period t is the sequence of derivatives $\partial C_{t+k}/\partial z_t^A$ that correspond to the elements $\beta_k(1,3)$ for $k = 0,\ldots,K$.

As described in the body of the paper, we use a two-step procedure to obtain the dynamic responses of compensation to unexpected firm performance. In the first step, we use SUR to obtain the structural coefficients (i.e., the Γ_ℓ matrices). In the second step, we invert the estimated structural system to obtain the moving-average form (i.e., the β_k matrices). We calculate these β_k matrices as follows:

$$\begin{aligned}
\beta_0 &= \Gamma_0^{-1}\\
\beta_1 &= (\Gamma_0^{-1}\Gamma_1)\Gamma_0^{-1}\\
\beta_2 &= [(\Gamma_0^{-1}\Gamma_1)(\Gamma_0^{-1}\Gamma_1) + (\Gamma_0^{-1}\Gamma_2)]\Gamma_0^{-1}\\
\beta_3 &= [\{(\Gamma_0^{-1}\Gamma_1)(\Gamma_0^{-1}\Gamma_1) + (\Gamma_0^{-1}\Gamma_2)\}(\Gamma_0^{-1}\Gamma_1) + (\Gamma_0^{-1}\Gamma_1)(\Gamma_0^{-1}\Gamma_2) + (\Gamma_0^{-1}\Gamma_3)]\Gamma_0^{-1}\\
&\vdots
\end{aligned}$$

The matrices of dynamic response follow $\beta_k = Q_k\Gamma_0^{-1}$, where $Q_k = 0$ for $k < 0$, $Q_0 = I$, and $Q_k = Q_{k-1}(\Gamma_0^{-1}\Gamma_1) + Q_{k-2}(\Gamma_0^{-1}\Gamma_2) + Q_{k-3}(\Gamma_0^{-1}\Gamma_3)$ for $k > 0$.

REFERENCES

Abowd, J. M. 1990. Does performance-based managerial compensation affect corporate performance? *Industrial and Labor Relations Review* 43 (February): 52–73.

Adut, D., W. Cready, and T. J. Lopez. 2003. Restructuring charges and CEO cash compensation. *The Accounting Review* 78 (1): 169–192.

Anderson, M. C., R. D. Banker, and S. Ravindran. 2000. Executive compensation in the information technology industry. *Management Science* 46 (April): 530–547.

Antle, R., and A. Smith. 1985. Measuring executive compensation: Methods and an application. *Journal of Accounting Research* 23 (Spring): 296–325.

———, and ———. 1986. An empirical investigation of the relative performance evaluation of corporate executives. *Journal of Accounting Research* 24 (Spring): 1–39.

Baber, W. R., S. Kang, and K. Kumar. 1998. Accounting earnings and executive compensation: The role of earnings persistence. *Journal of Accounting and Economics* 25 (May): 169–193.

———, ———, and ———. 1999. The explanatory power of earnings levels vs. earnings changes in the context of executive compensation. *The Accounting Review* 74 (October): 459–472.

Banker, R., and S. M. Datar. 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* 27 (Spring): 21–39.

Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (May/June): 637–654.

Boschen, J. F., and K. J. Smith. 1995. You can pay me now *and* you can pay me later: The dynamic response of executive compensation to firm performance. *Journal of Business* 68 (October): 577–608.

Bushman, R. M., and R. J. Indjejikian. 1993. Accounting income, stock price, and managerial compensation. *Journal of Accounting and Economics* 16 (January/April/July): 1–23.

———, E. Engel, J. Milliron, and A. Smith. 2000. An analysis of the relation between stewardship and valuation roles of earnings. Working paper, University of Chicago.

———, and A. Smith. 2001. Financial accounting information and corporate governance. *Journal of Accounting and Economics* 32 (December): 237–333.

DeFond, M. L., and C. W. Park. 1999. The effect of competition on CEO turnover. *Journal of Accounting and Economics* 27 (February): 35–56.

Dickey, D. A., and W. A. Fuller. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49 (July): 1057–1071.

Fama, E. F. 1980. Agency problems and the theory of the firm. *Journal of Political Economy* 88 (April): 288–307.

Feltham, G. S., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review* 69 (July): 429–453.

Gibbons, R., and K. J. Murphy. 1992. Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy* 100 (June): 468–505.

Granger, C. W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37 (August): 424–438.

Hall, B. J., and J. B. Liebman. 1998. Are CEOs really paid like bureaucrats? *Quarterly Journal of Economics* 113 (August): 653–691.

Hamilton, J. D. 1994. *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hayes, R. M., and S. Schaefer. 2000. Implicit contracts and the explanatory power of top executive compensation for future performance. *The Rand Journal of Economics* 31 (Summer): 273–293.

Indjejikian, R. J., and D. Nanda. 1999. Dynamic incentives and responsibility accounting. *Journal of Accounting and Economics* 27 (April): 177–201.

Jensen, M., and K. J. Murphy. 1990. Performance pay and top-management incentives. *Journal of Political Economy* 98 (April): 225–264.

Lambert, R. A. 1983. Long-term contracts and moral hazard. *Bell Journal of Economics* 14 (Autumn): 441–452.

———, and D. F. Larcker. 1987. Analysis of the use of accounting and market measures in executive compensation contracts. *Journal of Accounting Research* 25 (Supplement): 85–129.

———. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32 (December): 3–87.

Masson, R. 1971. Executive motivations, earnings, and consequent equity performance. *Journal of Political Economy* 79 (November/December): 1278–1292.

Murphy, K. 1985. Corporate performance and managerial remuneration. *Journal of Accounting and Economics* 7 (April): 11–42.

———. 1986. Incentives, learning, and compensation: A theoretical and empirical investigation of managerial labor contracts. *Rand Journal of Economics* 17 (Spring): 59–76.

———, and J. L. Zimmerman. 1993. Financial performance surrounding CEO turnover. *Journal of Accounting and Economics* 16 (January-July): 273–315.

———. 1999. Executive compensation. In *Handbook of Labor Economics*, Vol. 3b, edited by O. Ashenfelter, and D. Card, Chapter 38, 2485–2563. New York, NY: North Holland.

———. 2000. Performance standards in incentive contracts. *Journal of Accounting and Economics* 30 (December): 245–278.

Rogerson, W. P. 1985. Repeated moral hazard. *Econometrica* 53 (January): 69–76.

Rosen, S. 1992. Contracts and the market for executives. In *Contract Economics*, edited by L. Werin, and H. Wijkander, 181–211. Oxford, U.K.: Basil Blackwell.

Sloan, R. 1993. Accounting earnings and top executive compensation. *Journal of Accounting and Economics* 16 (January/April/July): 55–100.

Warner, J. B., R. L. Watts, and K. H. Wruck. 1988. Stock prices and tcp management changes. *Journal of Financial Economics* 20 (January/March): 461–492.

Weisbach, M. S. 1988. Outside directors and CEO turnover. *Journal of Financial Economics* 20 (January/March): 431–460.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 169–192

Restructuring Charges and CEO Cash Compensation: A Reexamination

Davit Adut
Texas A&M University
William H. Cready
Louisiana State University
Thomas J. Lopez
Georgia State University

ABSTRACT: Prior research generally concludes that compensation committees completely shield executive compensation from the effect of restructuring charges on earnings. In contrast, we find that after controlling for the growth in annual inflation-adjusted CEO cash compensation, compensation committees only partially shield CEO compensation from the adverse effect of restructuring charges on earnings, on average. In further analyses, we identify factors associated with cross-sectional differences in the extent of shielding. Specifically, we find that compensation committees appear to: (1) *completely shield* initial and subsequent restructuring charges for CEOs with long tenure, provided that the firm had not recorded a charge in the two immediately prior years; (2) provide *no shielding* of subsequent restructuring charges taken by short-tenured CEOs if the firm reported a prior restructuring charge within two years of the current charge; (3) and *partially shield* the other categories of restructuring charges. Overall, this study provides evidence that compensation committees evaluate the context of each restructuring in determining the extent to which they will intervene to shield executive compensation from the effect of these charges.

Keywords: *executive compensation; restructuring charges; compensation committee.*

We thank Larry Brown, Bill Baber, Michael Clement, Lynn Rees, K. Sivaramakrishnan, Mike Kinney, Mike Wilkins, Ed Swanson, Mary Ellen Carter, and John Core (our discussant at the 2001 AAA Annual Meeting) for their valuable contributions. We also gratefully acknowledge the suggestions of workshop participants at Georgia State University, The George Washington University, The University of Texas at Dallas, Texas A&M University, Louisiana State University, and the 2001 American Accounting Association Annual meeting in Atlanta, GA. We especially thank Professor Mark Huson for providing us with the listing of firms included in the Dechow et al. (1994) analysis, and two anonymous reviewers.

Submitted April 2001
Accepted September 2002

Data Availability: *A list of the firms is available from the authors. All other data are available from public sources identified in the text.*

I. INTRODUCTION

Compensation committees administer corporate executive annual incentive compensation plans. These committees decide whether and how to modify GAAP-based accounting income in determining executive compensation. To obtain insights into the role accounting information plays in this compensation decision, we investigate the degree to which compensation committees assess the context of a specific income item—restructuring charges—when deciding whether to shield CEO compensation from the adverse effects of the restructuring charge on current period income. For example, do prior restructuring charges, the length of the CEO's tenure, and the magnitude of the charge influence these compensation decisions?

Dechow et al. (1994) (hereafter DHS) make an important contribution to the literature by providing the first empirical evidence that compensation committees intervene to modify GAAP-based income in determining executive compensation. DHS also provide evidence suggesting that this intervention depends on certain contextual factors. For example, they show that compensation committees are less likely to shield executive compensation when the firm is a frequent restructurer, and that the longer the CEO has been in office, the more likely the compensation committee will shield his compensation from the effect of restructuring charges.

We extend DHS's analysis in two respects. First, we reexamine their specific conclusion that, on average, compensation committees fully shield chief executive officers' (CEOs') cash-based compensation from the earnings effect of restructuring charges.[1] (In fact, their empirical evidence is consistent with CEOs being awarded additional bonus pay for undertaking restructurings.) Second, we also extend DHS's contextual factor analysis by providing finer insights into how the frequency of reported restructuring charges, temporal proximity of prior restructuring charges, and CEO tenure influence whether compensation committees provide full, partial, or no shielding of CEO cash compensation from restructuring charges. Hence, our analysis provides insights into how these factors affect CEO compensation.

Our results confirm DHS's main conclusions that compensation committees intervene to modify the income number used to determine CEO compensation, and that this intervention depends on the tenure of the CEO and the number of restructurings the firm has reported. However, our detailed analyses differ from DHS's more specific conclusion that compensation committees fully shield executive compensation from the effect of restructuring charges. Rather, our analysis suggests that compensation committees, on average, only partially shield CEO compensation from the negative effect of restructuring charges on current period earnings. The primary explanation for this difference between the two studies' results is that we control for the annual inflation-adjusted growth in executive compensation. DHS does control for inflation by using Consumer Price Index (CPI) deflators. However, because CEO cash compensation growth has routinely exceeded the inflation rate since the late 1960s, firm-specific time-series models of executive compensation should control for this temporal trend in executive compensation.

---

[1] *Shielding* is the degree to which compensation committees add the restructuring charge back to income before setting earnings-based bonuses. *Full shielding* implies that they add 100 percent of the charge back to income, while *no shielding* implies that they add none of the charge back. *Partial shielding* implies that they add part of the restructuring charge back to income before setting executive earnings-based bonuses.

Given the magnitude of this compensation trend, we reexamine how restructuring charge frequency and CEO tenure affect the extent to which compensation committees shield CEO compensation from the adverse effects of restructuring charges on earnings. Our analysis shows that the degree of shielding depends not only on the frequency of reported restructurings, but also on their historical pattern. Specifically, we find that compensation committees fully shield CEO compensation from the effect on earnings of the firm's initial restructuring charge. The treatment of subsequent charges, however, depends on their temporal proximity to any previously reported restructuring charges. If the restructuring charge occurs within two years of a prior charge (a temporally proximate subsequent charge), then our results suggest that compensation committees do not shield CEO compensation from the charge. If, however, more than two years have elapsed since the prior charge (a temporally nonproximate subsequent charge), then our results indicate that compensation committees partially shield CEO compensation from the restructuring charge.

Our results are generally consistent with DHS's evidence that the longer the executive's tenure, the more compensation committees shield executive compensation. However, even after controlling for CEO tenure, we find that the degree of shielding depends on whether the restructuring is an initial charge, a temporally nonproximate subsequent charge, or a temporally proximate subsequent charge. That is, our evidence suggests that the degree of shielding is a joint function of the length of CEO tenure and the historical pattern of the firm's reported restructurings. Specifically, we find that compensation committees appear to (1) completely shield initial and temporally nonproximate restructuring charges for CEOs with long tenure, (2) provide no shielding of temporally proximate subsequent charges taken by short-tenured CEOs, and (3) partially shield the other categories of restructuring charges.

In summary, prior research generally concludes that compensation committees completely shield executive compensation from the adverse effect on current period earnings of potentially value-enhancing restructurings (DHS; Gaver and Gaver 1998). Our results, however, suggest that compensation committees do not view all restructurings favorably. This is consistent with prior research's conclusions and anecdotal evidence that restructurings are not always value-enhancing (Brickley and Van Drunen 1990; Carter 2000; Atiase et al. 2001; Lopez et al. 2002), and that some managers appear to use restructuring charges to manage earnings from one period to another (Moehrle 2002; Levitt 1998). Our evidence suggests that compensation committees evaluate the context of individual restructurings in determining the degree to which they will intervene to protect executive compensation from the effect of these charges. We interpret our evidence as suggesting that compensation committees assess a compensation penalty on the CEO in contexts where the restructuring charge is more likely to be opportunistic.

We have organized the remainder of the paper as follows. Section II develops the research design. Section III presents the sample selection and descriptive statistics. Section IV presents empirical results, and Section V briefly summarizes and concludes.

II. RESEARCH DESIGN

Base Model

DHS use the following firm-specific time-series model of CEO compensation:

$$COMP_t = \beta_0 + \beta_1 ADJ\_INC_t + \beta_2 R\_CHARGE_t + \epsilon_t \quad (1)$$

where:

$COMP_t$ = the CEO's cash compensation (salary and bonus) in year t, in inflation-adjusted dollars;

R_CHARGE_t = the restructuring charge in year t, in inflation-adjusted dollars; and
ADJ_INC_t = pre-restructuring charge income, computed as $INCOME_t$ − R_CHARGE_t, where $INCOME_t$ equals earnings before tax, extraordinary items, and the results of discontinued operations in year t, in inflation-adjusted dollars.

Like DHS, we define CEO compensation (COMP) as base salary plus bonus, and exclude other components of executive compensation, such as stock options. We focus on CEO cash compensation to facilitate comparison with DHS and Gaver and Gaver (1998), and because the bonus component of cash compensation typically requires achievement of a target annual earnings mark.[2] We inflation-adjust all amounts in our analyses to 1989 dollars based on the CPI published by the Bureau of Labor Statistics.

Although we retain the DHS focus on cash compensation, we incorporate two additional determinants of executive compensation. First, Gaver and Gaver (1998) establish that the effect of one dollar of positive income exceeds the effect of one dollar of loss on CEO cash compensation. DHS address this by eliminating loss-year observations in a robustness test. In contrast, we follow Gaver and Gaver (1998) and estimate a separate loss-year earnings coefficient.

Second, executive compensation has increased faster than the rate of inflation (Byrne 1991; Muckian 1993; Pratt 1996; Romani 1997). The ratio of executive compensation to that of hourly wage earners is a common benchmark. For instance, data reported in Smart (1999) imply that, at a minimum, the ratio of average CEO cash compensation to average factory worker pay nearly doubles from 42 in 1980 to 81 in 1998.[3] Murphy (1999) reports that the average annual growth in inflation-adjusted CEO cash compensation is 2.6 percent from 1970 to 1996 and 3.5 percent from 1982 to 1996. Byrne (1991) suggests that earnings growth can explain only a portion of this growth in compensation.

In firm-specific time-series models of CEO compensation, failure to allow for this trend may mask the true effect of a determinant of CEO compensation that is correlated with time. In our case, we estimate firm-specific time-series compensation models over an extended period characterized by a marked increase in (real) CEO compensation, 1970–1997. Moreover, the compensation determinant in which we are most interested—restructuring charges—also increases over time. Specifically, no restructuring charges occur in the sample until 1982, and the median charge (in time) does not occur until 1992.[4] Thus, it is important to control for this upward trend in CEO compensation to assess the degree to which compensation committees shield executive compensation from the effect of restructuring charges.

We control for the trend in inflation-adjusted CEO cash compensation by developing an out-of-sample index. We randomly select 100 nonrestructuring firms in the *Forbes* annual compensation survey for each year in our time-series (1970–1997). From these randomly selected firms, we record the out-of-sample median CEO cash compensation for each year in our time-series. We then use the out-of-sample medians to measure the annual growth in inflation-adjusted CEO cash compensation (TREND) as the median inflation-adjusted CEO cash compensation in year t divided by the median inflation-adjusted CEO cash compensation in 1970.[5]

---

[2] During the time period we examine (1982–1997), cash compensation is, on average, approximately 72 percent of total compensation (Murphy 1999). However, by the end of our study period, equity-based compensation had increased to approximately 50 percent of total CEO compensation.

[3] The measure of compensation in Smart (1999) is total CEO compensation. However, Smart (1999) also indicates that in 1998, cash compensation amounts to only 20 percent of total executive compensation. Using that 20 percent estimate, we estimate that, at a minimum, the ratio of average CEO cash compensation to average factory worker pay rose from 42 in 1980 to 81 in 1998.

[4] The untabulated Pearson correlation between year and restructuring charge frequency is 0.87 (significant at the 0.0001 level). The correlation between year and mean inflation-adjusted CEO compensation is 0.92 (significant at the 0.0001 level).

Thus, we modify Equation (1) to form our base model as follows:

$$COMP_t = \alpha_0 + \alpha_1 POS\_INC_t + \alpha_2 NEG\_INC_t + \alpha_3 R\_CHARGE_t + \alpha_4 TREND_t + \epsilon_t \quad (2)$$

where:

POS_INC_t = ADJ_INC_t if ADJ_INC_t is positive, and 0 otherwise;
NEG_INC_t = ADJ_INC_t if ADJ_INC_t is negative, and 0 otherwise; and
$TREND_t$ = out-of-sample median inflation-adjusted compensation in year t divided by the out-of-sample median inflation-adjusted compensation in 1970.

Our estimates of Equation (2), like those of DHS, also include a first-order autocorrelation term.

We examine the degree of shielding using the DHS two-hypothesis framework. Defining α_1 as the (weighted average) mean of the firm-specific POS_INC coefficients and α_3 as the (weighted average) mean of the firm-specific R_CHARGE coefficients, if α_1 exceeds α_3 and α_3 is less than or equal to zero, then compensation committees completely shield CEO compensation from the effect of restructuring charges.[6] If α_1 exceeds α_3 and α_3 is greater than zero, then compensation committees partially shield CEO compensation from the effect of restructuring charges. Finally, if α_1 is less than or equal to α_3 and α_3 exceeds zero, then compensation committees do not shield executive compensation from the effect of restructuring charges.

Historical Pattern of Reported Restructurings Model

DHS report evidence suggesting that the more restructurings the firm undertakes, the less likely compensation committees are to shield CEO compensation. The DHS analysis, however, leaves several questions unresolved. First, because DHS find that the coefficient on restructuring charges is negative (i.e., that, on average, managers receive a bonus for undertaking a restructuring), it is not clear how to interpret evidence that compensation committees provide "less shielding" for repetitive restructurings. Thus, we directly assess whether compensation committees fully, partially, or do not shield CEO compensation from the earnings effect of repetitive restructuring charges. Second, DHS report the average effect of restructuring frequency on CEO compensation, per restructuring firm. However, our analysis probes deeper and investigates whether compensation committees treat restructuring charges differently within the same firm depending on the firm's historical pattern of the reported restructurings.

We expect compensation committees to shield initial restructuring charges more than subsequent restructuring charges. Probing further, we also expect the historical pattern of prior restructuring charges to affect the compensation committee's willingness to shield the CEO from the adverse effects of subsequent restructuring charges. Specifically, we expect compensation committees to view a subsequent charge after one in the immediately preceding two years (i.e., a temporally proximate subsequent charge) as less deserving of special treatment than a restructuring charge for which the prior charge occurred several years in the past (i.e., a temporally nonproximate subsequent charge). Thus, we separately identify the compensation effect of initial, temporally nonproximate subsequent, and temporally proximate subsequent charges, to evaluate how the historical pattern of the firm's restructuring charges affects compensation committees' propensity to shield CEO compensation from the effect of a contemporary restructuring charge.

---

[5] Our inferences are robust to several other measures of TREND. First, we use the same method to calculate TREND described above, except we use the median CEO cash compensation of the sample firms rather than out-of-sample medians. Second, we use the year-to-year percentage change in median CEO cash compensation, using both the out-of-sample medians and the sample medians. Last, we use year (70, 71,...,97) and the square root of year as measures of TREND.

[6] We weight the coefficients in inverse proportion to their standard errors, consistent with the Z-statistic in DHS. Appendix A describes this weighting process and documents its linkage to the Z-statistic.

We examine multiple restructuring charge firms by estimating the following expanded version of Equation (2):

$$\begin{aligned} COMP_t = {} & \zeta_0 + \zeta_1 POS\_INC_t + \zeta_2 NEG\_INC_t + \zeta_3 TREND_t + \zeta_4 INITIAL\_RC_t \\ & + \zeta_5 SUB\_NONPROX\_RC_t + \zeta_6 SUB\_PROX\_RC_t + \epsilon_t \end{aligned} \tag{3}$$

where:

INITIAL_RC = R_CHARGE$_t$ if the restructuring charge is the first the firm reported during the test period (1982–1997), otherwise 0;[7]

SUB_NONPROX_RC = R_CHARGE$_t$ if the restructuring charge is subsequent to the initial charge the firm reported during the test period (1982–1997), *and* the firm did not report a previous restructuring charge in the two years prior to year t, otherwise 0; and

SUB_PROX_RC = R_CHARGE$_t$ if the restructuring charge is subsequent to the initial charge the firm reported during the test period (1982–1997), *and* the firm reported a previous restructuring charge in the two years prior to year t, otherwise 0.

Equation (3) allows us to estimate directly the mean compensation effect of each type of charge (initial, temporally nonproximate subsequent, and temporally proximate subsequent) for the sample of multiple restructuring firms. We then repeat DHS-based tests for each category. That is, we test for partial shielding with respect to each type of restructuring charge by evaluating whether the mean coefficient for each type of charge exceeds zero. We further test whether the compensation effect differs among the three categories of charges by testing for differences among the coefficients for the three different restructuring charge classifications. Evidence that the coefficient on temporally nonproximate subsequent charges (ζ_5) or temporally proximate subsequent charges (ζ_6) is greater than the coefficient on initial charges (ζ_4) would support our contention that the historical pattern of reported restructurings affects the degree to which compensation committees shield CEO compensation from contemporaneous charges. That is, the frequency of reported restructuring charges and the temporal proximity of previously reported restructuring charges decrease the degree to which compensation committees shield CEO compensation from contemporary charges. A higher coefficient for temporally proximate (ζ_6) relative to temporally nonproximate (ζ_5) subsequent charges would indicate that recent prior restructuring charges further reduce the extent to which compensation committees shield executive compensation from current restructuring charges.

---

[7] Our design assumes that the initial charge (INITIAL_RC) in our data set (1982–1997) actually represents the initial restructuring charge the firm reported. To assess the validity of this assumption, we examined each sample firm's financial statements for the three years prior to 1982. That examination revealed only one restructuring charge reported by our sample firms for that three-year period (RCA, 1981: $230 million charge).

CEO Tenure Model

DHS posit that the longer the executive's expected horizon with the firm, the more likely the compensation committee will shield him from the effect of restructuring charges. Consistent with this argument, they show that the coefficient on restructuring charges is negatively associated with the number of years the executive has been in office. However, the DHS design does not indicate whether compensation committees fully, partially, or do not shield firms that have CEOs with long vs. short tenure. We therefore extend the DHS analysis by modifying Equation (2) as follows:

$$\begin{aligned}\mathrm{COMP}_t = {} & \theta_0 + \theta_1\mathrm{POS\_INC}_t + \theta_2\mathrm{NEG\_INC}_t + \theta_3\mathrm{TREND}_t + \theta_4\mathrm{LTENURE\_RC}_t \\ & + \theta_5\mathrm{STENURE\_RC}_t + \epsilon_t \end{aligned} \tag{4}$$

where:

$\mathrm{LTENURE\_RC}_t$ = $\mathrm{R\_CHARGE}_t$ if the charge is taken in year t by a CEO whose tenure in year t is at or above the median tenure of all CEOs who have reported a restructuring charge in the sample, otherwise 0; and

$\mathrm{STENURE\_RC}_t$ = $\mathrm{R\_CHARGE}_t$ if the charge is taken in year t by a CEO whose tenure in year t is below the median tenure of all CEOs who have reported a restructuring charge in the sample, otherwise 0.

We define the tenure of the CEO as *long tenure* (LTENURE_RC) if the number of years the executive has been the firm's CEO (as of the year of restructuring) is at or above the median tenure (seven years) of all CEOs in the sample in the year the firm records a restructuring charge. Similarly, a CEO has *short tenure* (STENURE_RC) if the number of years he has been the firm's CEO (as of the year of restructuring) is below the median tenure of all CEOs in the sample in the year the firm records a restructuring charge. Evidence that the coefficient on LTENURE_RC (θ_4) is smaller than the coefficient on STENURE_RC (θ_5) would suggest that compensation committees shield compensation of CEOs with longer tenure to a greater degree than compensation of CEOs with shorter tenure.

III. SAMPLE SELECTION AND DESCRIPTIVE STATISTICS

The full sample consists of firms that took at least one restructuring charge between 1982 and 1997. Restructuring charges arose along with the restructuring of the U.S. manufacturing industry in the 1980s, and DHS (1994, 149) indicate that *Accounting Trends and Techniques* did not even disclose restructuring charges in the list of disclosed losses prior to 1982. Thus, consistent with DHS, we start our search for restructuring charges in 1982. We received initial restructuring observations in our sample for the years prior to 1990 from Professor Mark Huson. For the years 1990 to 1994, we identified restructuring charges by searching the National Automated Accounting Research System (NAARS) using the search strings "restruct!" "unusual," and "special." For years after 1994, we search the Lexis-Nexis Academic Universe Business News database using the same search strings. To facilitate the collection of compensation data, we exclude firms identified in this search that are not among the *Fortune* 500 firms in 1992.

We estimate the compensation-earnings relation on a firm-by-firm basis using all (annual) compensation and earnings data available for the period from 1970 through 1997.[8] Hence, we have a maximum of 28 time-series observations for each firm. We require that the firm have a minimum of 15 years of compensation data to remain in the test sample. In addition, the firm must report at least one restructuring charge during its compensation time-series and have the data on Compustat necessary to estimate all equations.

Forbes reports salary and bonus data for CEOs of firms in any of their top 500 listings for the years 1970 through 1997.[9] When *Forbes* does not report the compensation of a CEO for a particular year, we manually collect the CEOs' cash compensation from the appropriate proxy statement.[10] We obtain financial statement data from the 1997 Compustat annual industrial files. Compustat does not separately report restructuring charges; therefore, we collect the dollar amounts for all restructuring charges directly from the annual income statement.

These data requirements result in a final sample of 129 firms reporting a total of 426 restructurings.[11] The mean and median years of compensation data for our sample are 26 and 28, respectively.

Table 1, Panel A reports that the number of restructuring charges ranges from a low of 4 in 1983 to a high of 44 in 1992. Panel A also reveals that our sample firms' inflation-adjusted CEO cash compensation increased dramatically over time. The mean inflation-adjusted cash compensation in 1982 was $653 thousand, whereas by 1997 the figure had more than doubled to over $1.4 million, an increase of 118 percent. The growth in inflation-adjusted CEO cash compensation is consistent with Murphy's (1999) evidence and with anecdotal media reports (Romani 1997; Pratt 1996; Muckian 1993; Byrne 1991). This dramatic growth highlights the importance of controlling for the trend in inflation-adjusted executive compensation when estimating time-series executive compensation models.

Panel B of Table 1 details the number of restructuring charges reported per firm. Of the 129 firms in the sample, only 27 firms (approximately 21 percent of the sample) report a single restructuring. More than 60 percent of the sample (78 firms) report three or more restructurings during the 16-year period from 1982 to 1997.

Table 2 reports descriptive statistics for the years in which sample firms report a restructuring charge. We use the CPI to inflation-adjust all amounts to 1989 dollars in this table. In the year of restructuring, our sample CEOs have been in office an average (median) of 8.65 (7.0) years. The sample generally consists of large firms, with mean total assets, market value of equity, and CEO cash compensation averaging $8.7 billion, $5.5 billion,

---

[8] To establish that our results are not significantly affected by including compensation and earnings data from the 1970 to 1981 time period, we reestimate all models using compensation and earnings data limited to the years 1982 through 1997 (maximum of 16 observations per firm). The results from these tests (not reported) are quantitatively and qualitatively similar to those reported in the paper.

[9] We limit our identification of restructuring firms to a search of the *Fortune* 500, as reported by NAARS, in order to facilitate the collection of CEO compensation data from the compensation survey of the *Forbes* 500. NAARS identifies firms that are listed in the *Fortune* 500, but not the *Forbes* 500. However, firms listed on the *Fortune* 500 and *Forbes* 500 consist of large public corporations.

[10] We collect compensation data from proxy statements if there are less than six years of missing compensation data from *Forbes* needed to complete the minimum of 15 years required. If more than five years of compensation data are missing, we exclude the firm from the sample.

[11] Our final sample excludes eight outlier firms identified by estimating Equation (2) on the 137 sample firms with complete data. Based on this analysis, we eliminated eight firms with a restructuring charge coefficient, R_CHARGE, that is three standard deviations greater than or less than the median restructuring charge coefficient of all firms in the sample. We reran all tests with the 137 firms that have complete data. The results of those tests (not reported) are quantitatively and qualitatively similar to those reported.

TABLE 1
Distribution of Restructuring Charges over Time and by Firm, and Mean Inflation-Adjusted CEO Cash Compensation over Time

Panel A: Distribution of Restructuring Charges and Mean Inflation-Adjusted CEO Cash Compensation by Fiscal Year

Year	*Number of Restructuring Charges*	*Mean Inflation-Adjusted CEO Cash Compensation ($Thousands)*[a]
1982	6	653
1983	4	739
1984	21	791
1985	30	804
1986	35	933
1987	30	922
1988	24	996
1989	19	1,030
1990	20	1,043
1991	34	967
1992	44	1,050
1993	41	1,055
1994	24	1,220
1995	34	1,251
1996	30	1,357
1997	30	1,425

Panel B: Number of Restructuring Charges Reported per Firm over the 1982–1997 Period

Restructuring Charges per Firm	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Total*
Number of firms	27	24	32	15	12	9	4	3	1	2	129

---

[a] We adjust all dollar amounts to 1989 equivalent dollars based on the CPI.

TABLE 2
Descriptive Statistics on 426 Restructuring Charges Reported between 1982 and 1997[a]

	Mean	*Standard Deviation*	*Lower Quartile*	*Median*	*Upper Quartile*
COMP ($THOUSAND)	996	644	604	872	1,201
TENURE (YEARS)	8.65	7.33	3	7	12
ASSETS ($MILLION)	8,654	20,743	1,685	3,256	7,923
MARKET VALUE ($MILLION)	5,488	9,068	1,065	2,384	5,983
R_CHARGE ($MILLION)	−243	464	−213	−78	−27
INCOME ($MILLION)	301	954	−14	121	424
ADJ_INC ($MILLION)	549	989	75	232	680
R_CHARGE/ASSETS	−0.048	0.104	−0.051	−0.028	−0.009
R_CHARGE/MARKET VALUE	−0.121	0.438	−0.079	−0.031	−0.012
R_CHARGE/ADJ_INC[b]	−0.823	2.913	−0.675	−0.248	−0.114

[a] We adjust all dollar amounts to 1989 equivalent dollars based on the CPI.

[b] Descriptive statistics on R_CHARGE/ADJ_INC excludes 61 restructurings where ADJ_INC is negative in the period in which the firm records a restructuring charge. These 61 restructuring observations were taken by 35 firms.

COMP = CEO cash compensation (salary and bonus) in inflation-adjusted dollars;
TENURE = the number of years the CEO has been in office in the year of the restructuring charge in year t;
ASSETS = total assets in year t, in inflation-adjusted dollars;
MARKET VALUE = market value of equity in year t, in inflation-adjusted dollars;
R_CHARGE = the restructuring charge in year t, in inflation-adjusted dollars;
INCOME = earnings before tax, extraordinary items, and the results of discontinued operations in year t, in inflation-adjusted dollars; and
ADJ_INC = INCOME − R_CHARGE.

and $996 thousand, respectively. The average restructuring charge of $243 million has a significant effect on the firm, averaging 4.8 percent of total assets, 12.1 percent of market value, and 82.3 percent of pre-restructuring charge income.[12]

IV. EMPIRICAL RESULTS

Empirical Tests of the Base Model: Equation (2)

Table 3 reports coefficient medians, weighted means, and Z-statistics from our 129 firm-specific estimations of Equation (2).[13] Consistent with DHS and Gaver and Gaver (1998), the weighted mean (median) POS_INC coefficient of 1.16 (0.98) is significantly positive at the 0.01 level. The weighted mean (median) coefficient on R_CHARGE of 0.64 (0.55) is also significantly positive, suggesting that compensation committees, at most, only partially shield CEO compensation from the effect of restructuring charges. We also find

---

[12] We do not compute the ratio of restructuring charge to pre-restructuring charge income for the 61 cases with negative pre-restructuring charge income.

[13] In addition to the parametric Z-statistic described in Appendix A, we also employ the nonparametric Wilcoxon signed rank test, a test of differences in location, to assess the significance of the coefficients and coefficient differences.

TABLE 3
Results from 129 Firm-Specific Regressions of CEO Cash Compensation on Income and Restructuring Charges, 1982–1997

Equation (2): $COMP_t = \alpha_0 + \alpha_1 POS\_INC_t + \alpha_2 NEG\_INC_t + \alpha_3 R\_CHARGE_t + \alpha_4 TREND_t$[a]

	Predicted Sign	*Median*[b]	*Weighted Mean Coefficient*	*Z-Statistic*[c]
POS_INC, α_1	+	0.98**	1.16	32.01**
NEG_INC, α_2 (n = 67)[d]	?	0.01	−0.17	−0.73
R_CHARGE, α_3	+	0.55**	0.64	5.09**
TREND, α_4	+	254.34**	278.94	70.38**
POS_INC − R_CHARGE, $\alpha_1 - \alpha_3$	+	0.32**	0.51	3.10**
Adjusted R^2		0.61		
Number of Observations		26		

---

*, ** Significant at 0.05 and 0.01, respectively.

[a] COMP = CEO cash compensation (salary and bonus) in inflation-adjusted dollars;
R_CHARGE = the restructuring charge in year t, in inflation-adjusted dollars;
ADJ_INC = INCOME − R_CHARGE, where INCOME equals earnings before tax, extraordinary items, and the results of discontinued operations in year t, in inflation-adjusted dollars;
POS_INC = ADJ_INC if ADJ_INC > 0;
NEG_INC = ADJ_INC if ADJ_INC < 0; and
TREND = inflation-adjusted CEO cash compensation growth index.

[b] The significance levels indicated in the Median column are based on the Wilcoxon signed rank test.

[c] Appendix A describes the weighted mean and Z-statistic calculations.

[d] Of the 129 firms in the sample, 67 firms experienced at least one loss in the sample time period (1970–1997). Accordingly, the mean and median for this coefficient are based on only 67 observations.

that the weighted mean (median) POS_INC coefficient is significantly higher than the coefficient on R_CHARGE (p-value < 0.01). That is, our results suggest that while compensation committees do place some weight on restructuring charges in determining CEO compensation, they place less weight on restructuring charges than on positive pre-restructuring charge income.[14] Hence, we conclude that compensation committees, on average, partially shield CEO compensation from the earnings effect of restructuring charges.[15]

---

[14] We focus on how the restructuring charge component of income affects CEO cash compensation, in years when the firm reports positive income, because Gaver and Gaver (1998, 245) establish that the relation between firm income and CEO cash compensation is essentially nullified in years when the firm reports a loss. However, we reran Equation (2) including a loss year dummy × restructuring charge interaction term to investigate whether compensation committees treat restructuring charges differently in years in which the firm reports a loss. The coefficient on the interaction term is not significantly different from zero. In other untabulated analyses, we find that the weighted mean (median) difference between the coefficients on NEG_INC and R_CHARGE is insignificant, estimated on the subsample of 35 firm-years in which both a loss and restructuring charge are recorded. This result suggests that compensation committees treat restructuring charges recorded in years when the firm reports a loss, in a manner consistent with the other components of losses.

[15] In additional tests, we partition the restructuring charge variable in Equation (2) based on the magnitude of the charge (large charge above the median and small charge below the median of all charges) and reestimate the regression. Results (not reported) suggest partial shielding for both large (above-median) and small (below-median) restructuring charges. We did not find a statistically significant difference between the coefficients on large and small charges.

Table 3 also reveals the importance of controlling for the temporal trend in inflation-adjusted CEO cash compensation. Consistent with our expectation, the weighted mean (median) coefficient on TREND is positive and highly significant (0.01 level). That is, our results suggest that the most important determinant of inflation-adjusted CEO cash compensation is the passage of time. Consequently, it is important to control for this temporal trend in firm-specific time-series compensation models.[16]

DHS conclude that, on average, compensation committees *completely shield* CEO compensation from restructuring charges. Their results further suggest that CEOs are awarded bonus pay that increases proportionately with the magnitude of the restructuring charge. The evidence we report in Table 3 is inconsistent with that conclusion. The primary difference between our results and the DHS results is that DHS report a significantly negative coefficient on R_CHARGE, whereas we report a significantly positive coefficient on R_CHARGE. To isolate the reason for this difference, we replicate the DHS analysis in Appendix B. This replication reveals that the primary explanation for the difference between our results and DHS stems from our control for TREND, which Table 3 shows is the most significant variable in our firm-specific CEO compensation model. Thus, the specific DHS conclusion that compensation committees *completely* shield CEO compensation from the effects of restructuring charges appears to be a spurious result of a correlated omitted variable—the dramatic increase in inflation-adjusted CEO cash compensation over time.

Empirical Tests of the Historical Pattern Model: Equation (3)

Our empirical tests of Equation (3) include only the subsample of 90 firms reporting more than one restructuring charge and at least one temporally proximate subsequent restructuring charge. These 90 firms recorded a total of 371 restructuring charges over the sample time period. The mean initial (INITIAL_RC), temporally nonproximate subsequent (SUB_NONPROX_RC), and temporally proximate subsequent (SUBSEQ_PROX_RC) restructuring charges are $191, $290, and $180 million, respectively. Most of the initial charges (57 out of 90) occur in the pre-1990 time period, whereas most of the temporally proximate subsequent (151 out of 228) and temporally nonproximate subsequent (50 out of 57) charges occur post-1989. However, in an untabulated supplemental analysis, we find no significant evidence of a between-decade shift in the restructuring charge coefficient for any of the three types of charges.

Panel A of Table 4 reports estimates for the Equation (3) specification. The weighted mean (median) coefficient for INITIAL_RC is insignificant at conventional levels. Thus, our evidence suggests that compensation committees fully shield CEO cash compensation from the adverse effects of initial restructuring charges.[17] In contrast, the weighted mean (median) coefficients for both types of subsequent charges, SUB_NONPROX_RC and SUB_PROX_RC, are positive and significant at the 0.01 level, indicating that compensation committees do not fully shield subsequent charges.

Panel B of Table 4 reports results for tests of differences among the Equation (3) coefficient estimates. Consistent with the Panel A inference that compensation committees

---

[16] Prior research establishes that CEO compensation increases with firm size (Bliss and Rosen 2001). Thus, we reestimate all models including firm size (log of total assets) as an additional independent variable. The results of these tests (not reported) are qualitatively and quantitatively similar to the results reported throughout the paper.

[17] In an untabulated analysis, we estimate the mean restructuring charge coefficient for the 27 firms that experience only one restructuring charge over the sample period. The weighted mean R_CHARGE coefficient is 0.20 and the associated Z-statistic is 0.49, suggesting that compensation committees fully shield CEO compensation from the adverse effects of a single restructuring charge. Hence, compensation committee treatment of these one-time charges appears indistinguishable from the treatment accorded initial charges by multiple restructuring charge firms.

TABLE 4

Results from 90 Firm-Specific Regressions of CEO Cash Compensation on Income and Initial, Temporally Nonproximate Subsequent, and Temporally Proximate Subsequent Restructuring Charges

Equation (3): $COMP_t = \zeta_0 + \zeta_1 POS\_INC_t + \zeta_2 NEG\_INC_t + \zeta_3 TREND_t + \zeta_4 INITIAL\_RC_t + \zeta_5 SUB\_NONPROX\_RC_t + \zeta_6 SUB\_PROX\_RC_t$[a]

Panel A: Coefficient Estimates

Coefficient	*n*[b]	*Median*[c]	*Weighted Mean Coefficient*	*Z-Statistic*[d]
POS_INC, ζ_1	90	1.02**	1.23	29.62**
NEG_INC, ζ_2	48	−0.01	−0.19	−0.60
TREND, ζ_3	90	284.15**	290.37	68.81**
INITIAL_RC, ζ_4	90	0.13	0.14	0.93
SUB_NONPROX_RC, ζ_5	52	0.35*	0.46	2.54**
SUB_PROX_RC, ζ_6	90	0.62**	0.90	3.95**
Adjusted R^2	90	0.57		

Panel B: Tests of Panel A Coefficient Differences

Coefficient Difference	*n*[b]	*Median*[c]	*Weighted Mean Difference*	*Z-Statistic*[d]
POS_INC − INITIAL_RC, $\zeta_1 - \zeta_4$	90	0.87**	1.09	6.30**
POS_INC − SUB_NONPROX_RC, $\zeta_1 - \zeta_5$	52	0.73**	0.61	2.88**
POS_INC − SUB_PROX_RC, $\zeta_1 - \zeta_6$	90	0.15	0.33	0.49
INITIAL_RC − SUB_NONPROX_RC, $\zeta_4 - \zeta_5$	52	−0.21	−0.25	−1.09
INITIAL_RC − SUB_PROX_RC, $\zeta_4 - \zeta_6$	90	−0.73**	−0.76	−3.67**
SUB_NONPROX_RC − SUB_PROX_RC, $\zeta_5 - \zeta_6$	52	−0.37**	−0.43	−2.50**

---

*, **Significant at 0.05 and 0.01, respectively.

[a] COMP = CEO cash compensation (salary and bonus) in inflation-adjusted dollars;
R_CHARGE = the restructuring charge in year t, in inflation-adjusted dollars;
INITIAL_RC = R_CHARGE_t if the restructuring charge is the first reported by the firm during the test period (1982–1997), otherwise 0;
SUB_NONPROX_RC = R_CHARGE_t if the restructuring charge is subsequent to the initial charge (INITIAL_RC) reported by the firm during the test period (1982–1997) and the firm did not report a previous restructuring charge in the two years prior to year t, otherwise 0;
SUB_PROX_RC = R_CHARGE_t if the restructuring charge is subsequent to the initial charge (INITIAL_RC) reported by the firm during the test period (1982–1997) and the firm reported a previous restructuring charge in the two years prior to year t, otherwise 0;
ADJ_INC = INCOME − R_CHARGE, where INCOME equals earnings before tax, extraordinary items, and the results of discontinued operations in year t, in inflation-adjusted dollars;
POS_INC = ADJ_INC if ADJ_INC > 0;
NEG_INC = ADJ_INC if ADJ_INC < 0; and
TREND = inflation-adjusted CEO cash compensation growth index.

[b] n is the number of firm-specific parameters used to compute the test statistics.

[c] The significance levels indicated in the Median column are based on the Wilcoxon signed rank test.

[d] Appendix A describes the weighted mean and Z-statistic calculations.

at least partly shield CEO cash compensation from the effects of restructuring charges, the weighted mean (median) POS_INC coefficient is significantly larger than both the INITIAL_RC and SUB_NONPROX_RC coefficients at the 0.01 level. Taken together with the insignificant coefficient on INITIAL_RC and the significantly positive coefficient on SUB_NONPROX_RC, this evidence indicates that compensation committees fully shield CEO compensation from the earnings effect of initial restructuring charges, and partially shield CEO compensation from the earnings effect of temporally nonproximate restructuring charges. In contrast, the weighted mean (median) difference between the POS_INC and SUB_PROX_RC coefficients is statistically indistinguishable from zero. Hence, we find no evidence that compensation committees shield executive compensation from temporally proximate subsequent charges.

The tests of differences among the restructuring charge coefficients confirm that the weighted mean (median) SUB_PROX_RC coefficient is significantly larger than both the INITIAL_RC and the SUB_NONPROX_RC coefficients. In contrast, the difference between the SUB_NONPROX_RC and INITIAL_RC coefficients is not significant at conventional levels. Hence, our results suggest that compensation committees shield initial and nonproximate restructuring charges to a greater extent than proximate restructuring charges that occur shortly after a prior restructuring charge.

Empirical Tests of the CEO Tenure Model: Equation (4)

DHS report evidence suggesting that the longer the CEO's tenure with the firm, the more the compensation committee will shield the executive from the adverse effect of restructuring charges. Table 5 reports estimates of our Equation (4) specification that further probes whether compensation committees fully shield, partially shield, or do not shield compensation of CEOs with long vs. short tenure, after controlling for the growth in inflation-adjusted CEO cash compensation.

The weighted mean (median) coefficients for both LTENURE_RC and STENURE_RC are positive and significant at the 0.05 level or better, and POS_INC is significantly greater (p-value < 0.01) than both LTENURE_RC and STENURE_RC. Taken together, these results suggest that compensation committees partially shield CEO cash compensation from the earnings effect of restructuring charges, whether the CEO has been in place for a short or long period of time. Contrary to DHS and our own expectation, however, we find no significant difference between the weighted mean and median coefficients on LTENURE_RC and STENURE_RC, after we control for the growth in inflation-adjusted CEO cash compensation.[18]

Combined Historical Pattern and CEO Tenure Model

Table 4 reports evidence that the degree to which compensation committees shield executive compensation from the adverse effects of restructuring charges depends on the firm's historical pattern of reporting restructurings. Contrary to our expectations, however, Table 5 reports that the extent of shielding does not differ significantly between CEOs with short tenure vs. long tenure. However, it is possible that differences in the number and pattern of reported restructuring charges confounds our analysis of the effects of CEO tenure. To disentangle the effects of CEO tenure from the effects of the firm's historical

---

[18] Defining LTENURE_RC and STENURE_RC based on the CEOs in the upper quartile and lower quartile of the sample yields identical inferences. In addition, although prior research suggests that new CEOs engage in "big bath" behavior (Strong and Meyer 1987; Elliott and Shaw 1988; Pourciau 1993), in supplementary analyses we found no evidence that compensation committees shield executives in their first year in office to a greater extent than executives who have been with the firm for a longer time period.

TABLE 5

Results of 129 Firm-Specific Regressions of CEO Cash Compensation on Income and Restructuring Charges for CEOs with Long Tenure vs. Short Tenure

Equation (4): $COMP_t = \theta_0 + \theta_1 POS\_INC_t + \theta_2 NEG\_INC_t + \theta_3 TREND_t + \theta_4 LTENURE\_RC_t + \theta_5 STENURE\_RC_t$[a]

	n^b	Median[c]	Weighted Mean Coefficient	Z-Statistic[d]
POS_INC, θ_1	129	1.00**	1.18	33.75**
NEG_INC, θ_2	67	−0.03	−0.14	−0.52
TREND, θ_3	129	264.20**	284.55	71.39**
LTENURE_RC, θ_4	73	0.49*	0.40	2.47**
STENURE_RC, θ_5	108	0.69**	0.72	4.01**
POS_INC − LTENURE_RC, $\theta_1 - \theta_4$	73	0.55**	0.96	4.19**
POS_INC − STENURE_RC, $\theta_1 - \theta_5$	108	0.41**	0.45	2.63**
LTENURE_RC − STENURE_RC, $\theta_4 - \theta_5$	53	0.16	0.26	1.58
Adjusted R^2		0.58		

---

*, ** Significant at 0.05 and 0.01, respectively.

[a] COMP = CEO cash compensation (salary and bonus) in inflation-adjusted dollars;
R_CHARGE = the restructuring charge in year t, in inflation-adjusted dollars;
LTENURE_RCt = R_CHARGE, if the charge is taken in year t by a CEO whose tenure in year t is at or above the median tenure of all CEOs who have restructuring charges in the sample, otherwise 0;
STENURE_RCt = R_CHARGE, if the charge is taken in year t by a CEO whose tenure in year t is below the median tenure of all CEOs who have restructuring charges in the sample, otherwise 0;
ADJ_INC = INCOME − R_CHARGE, where INCOME equals earnings before tax, extraordinary items, and the results of discontinued operations in year t, in inflation-adjusted dollars;
POS_INC = ADJ_INC if ADJ_INC > 0;
NEG_INC = ADJ_INC if ADJ_INC < 0; and
TREND = inflation adjusted CEO cash compensation growth index.

[b] n is the number of firm-specific parameters used to compute the test statistics.

[c] The significance levels indicated in the Median column are based on the Wilcoxon signed rank test.

[d] Appendix A describes the weighted mean and Z-statistic calculations.

---

pattern of reported restructuring charges, we integrate our analysis of the historical pattern of reported restructurings into our analysis of the effect of CEO tenure as follows:

$$\begin{aligned} COMP_t = {} & \lambda_0 + \lambda_1 POS\_INC_t + \lambda_2 NEG\_INC_t + \lambda_3 TREND_t + \lambda_4 LT\_INITIAL\_RC_t \\ & + \lambda_5 LT\_SUB\_NONPROX\_RC_t + \lambda_6 LT\_SUB\_PROX\_RC_t \\ & + \lambda_7 ST\_INITIAL\_RC_t + \lambda_8 ST\_SUB\_NONPROX\_RC_t \\ & + \lambda_9 ST\_SUB\_PROX\_RC_t. \end{aligned} \tag{5}$$

The prefix LT_(ST_) indicates that the firm recorded a restructuring in a year when the CEO's tenure is at or above (below) the median tenure of the CEO in the restructuring charge year for our sample of 426 charges. All other variables are as previously identified.

Equation (5) allows us to directly estimate the compensation effect of each type of charge (initial, temporally nonproximate subsequent, and temporally proximate subsequent) for the sample of 90 multiple restructurers, conditional on the length of the CEO's tenure. Table 6 reports summary statistics for our firm-specific estimations of Equation (5). The weighted mean (median) coefficients on LT_INITIAL_RC and LT_SUB_NONPROX_RC are not significant at conventional levels, consistent with full shielding. On the other hand, the weighted mean (median) coefficients on the other restructuring charge classifications are all significantly positive at the 0.05 level or better. Thus, our evidence suggests that compensation committees fully shield CEOs with long tenure for all restructuring charge types except temporally proximate subsequent restructuring charges. The evidence further suggests that compensation committees, at best, only partially shield short-tenured CEOs' compensation from the effect of any of the three kinds of restructurings. Overall, the Table 6 results are consistent with compensation committees providing a benefit to long-tenured CEOs in the degree to which they shield executive compensation from occasional restructuring charges.

Panel A of Table 7 reports results for tests of differences between the POS_INC coefficient and the restructuring charge coefficients reported in Table 6. The weighted mean (median) POS_INC coefficient is significantly greater than the coefficients on long-tenured initial ($\lambda_1 - \lambda_4$), long-tenured temporally nonproximate ($\lambda_1 - \lambda_5$), long-tenured temporally proximate ($\lambda_1 - \lambda_6$), and short-tenured initial ($\lambda_1 - \lambda_7$) restructuring charges at the 0.05 level or better, indicating significant shielding of all restructuring charges taken by long-tenured CEOs and initial charges taken by short-tenured CEOs. The weighted mean (median) POS_INC coefficient is insignificantly different than the coefficients on short-tenured temporally nonproximate ($\lambda_1 - \lambda_8$) and short-tenured temporally proximate ($\lambda_1 - \lambda_9$) restructuring charges, suggesting that compensation committees provide little, if any, shielding of cash compensation from subsequent restructuring charges taken by short-tenured CEOs.

The short-tenured restructuring charge coefficients reported in Table 6 uniformly exceed their long-tenured counterparts, consistent with less shielding for short-tenured CEOs compared to long-tenured CEOs. However, Panel B of Table 7 reports that this difference is significant (at the 0.01 level) for only the temporally nonproximate subsequent ($\lambda_5 - \lambda_8$) and temporally proximate subsequent ($\lambda_6 - \lambda_9$) restructuring charge classifications.[19] These coefficient differences reveal that, in contrast to the insignificant results in Table 5, a more fully specified model supports DHS's conclusion that compensation committees provide more shielding for long-tenured CEOs—at least for those CEOs that take multiple restructuring charges.

Consistent with Table 4, for both the long-tenured and short-tenured CEOs, the weighted mean restructuring charge coefficient is smallest for initial restructuring charges and greatest for temporally proximate subsequent charges. Panel C of Table 7 shows, however, that these differences are not significant at conventional levels for the long-tenured CEOs. In contrast, for the short-tenured CEOs, the temporally proximate subsequent restructuring charge coefficient is significantly higher than the coefficients on both initial ($\lambda_7 - \lambda_9$) and temporally nonproximate subsequent ($\lambda_8 - \lambda_9$) restructuring charges, indicating that the frequency of reporting restructuring charges and the temporal proximity of

---

[19] The Z-test reported in Panels B and C of Table 7 differs slightly from the tests reported elsewhere in the paper. This portion of the analysis uses the two-sample form of the Z-statistic described in Appendix A, not the within-firm Z-test employed elsewhere in the paper. We use the two-sample form of the Z-test for this analysis, because most firms do not have initial, proximate, and nonproximate restructuring charges taken by a CEO with short tenure and (in other years) taken by CEOs with long tenure.

TABLE 6
Results from 90 Firm-Specific Regressions of CEO Cash Compensation on Income and Initial, Temporally Nonproximate Subsequent, and Temporally Proximate Subsequent Restructuring Charges, Conditional on Whether the Restructuring Charge is Reported by a CEO with Long vs. Short Tenure

Equation (5):

$$\begin{aligned} COMP_t = {} & \lambda_0 + \lambda_1 POS\_INC_t + \lambda_2 NEG\_INC_t + \lambda_3 TREND_t + \lambda_4 LT\_INITIAL\_RC_t \\ & + \lambda_5 LT\_SUB\_NONPROX\_RC_t + \lambda_6 LT\_SUB\_PROX\_RC_t \\ & + \lambda_7 ST\_INITIAL\_RC_t + \lambda_8 ST\_SUB\_NOPROX\_RC_t \\ & + \lambda_9 ST\_SUB\_PROX\_RC_t{}^{a} \end{aligned}$$

	n[b]	Median[c]	Weighted Mean Coefficient	Z-Statistic[d]
POS_INC, λ_1	90	1.04**	1.24	30.82**
NEG_INC, λ_2	48	−0.00	−0.26	−0.51
TREND, λ_3	90	265.33**	275.99	69.41**
LT_INITIAL_RC, λ_4	49	0.05	0.08	0.40
LT_SUB_NONPROX_RC, λ_5	25	0.25	0.23	0.89
LT_SUB_PROX_RC, λ_6	56	0.34**	0.39	2.63**
ST_INITIAL_RC, λ_7	48	0.54**	0.35	2.97**
ST_SUB_NONPROX_RC, λ_8	36	0.67*	0.60	3.04**
ST_SUB_PROX_RC, λ_9	53	0.89**	1.18	6.00**
Adjusted R^2	90	0.55		

---

*, ** Significant at 0.05 and 0.01, respectively.

[a] COMP = CEO cash compensation (salary and bonus) in inflation-adjusted dollars;
R_CHARGE = the restructuring charge in year t, in inflation-adjusted dollars;
LT_(ST_) = prefix indicates the restructuring charge variable is recorded in a year when the CEO's tenure is at or above (below) the median tenure of the CEO in the restructuring charge year for test sample of restructuring charges.
INITIAL_RC = R_CHARGE, if the restructuring charge is the first reported by the firm during the test period (1982–1997), otherwise 0;
SUB_NONPROX_RC = R_CHARGE, if the restructuring charge is subsequent to the initial charge (INITIAL_RC) reported by the firm during the test period (1982–1997) and the firm did not report a previous restructuring charge in the two years prior to year t, otherwise 0;
SUB_PROX_RC = R_CHARGE, if the restructuring charge is subsequent to the initial charge (INITIAL_RC) reported by the firm during the test period (1982–1997) and the firm reported a previous restructuring charge in the two years prior to year t, otherwise 0;
ADJ_INC = INCOME − R_CHARGE, where INCOME equals earnings before tax, extraordinary items, and the results of discontinued operations in year t, in inflation-adjusted dollars;
POS_INC = ADJ_INC if ADJ_INC > 0;
NEG_INC = ADJ_INC if ADJ_INC < 0; and
TREND = inflation-adjusted CEO cash compensation growth index.

[b] n is the number of firm-specific parameters that are used to compute the test statistics.

[c] The significance levels indicated in the Median column are based on the Wilcoxon signed rank test.

[d] Appendix A describes the weighted mean and Z-statistic calculations.

TABLE 7
Tests of Differences in Equation (5) Restructuring Charge Coefficients[a]

Panel A: Income Coefficient Differences

Coefficient Difference	*n*[b]	*Median*[c]	*Weighted Mean Difference*	*Z-Statistic*[d]
POS_INC − LT_INITIAL_RC, $\lambda_1 - \lambda_4$	49	0.94**	0.85	3.31**
POS_INC − LT_SUB_NONPROX_RC, $\lambda_1 - \lambda_5$	25	0.87*	0.92	1.95*
POS_INC − LT_SUB_PROX_RC, $\lambda_1 - \lambda_6$	56	0.31*	0.28	1.70
POS_INC − ST_INITIAL_RC, $\lambda_1 - \lambda_7$	48	0.30*	0.44	2.59**
POS_INC − ST_SUB_NONPROX_RC, $\lambda_1 - \lambda_8$	36	0.23	0.30	1.34
POS_INC − ST_SUB_PROX_RC, $\lambda_1 - \lambda_9$	53	0.10	0.07	0.49

Panel B: Differences by CEO Tenure

Coefficient Difference	*n*[b]	*Median*[c]	*Weighted Mean Difference*	*Z-Statistic*[d]
LT_INITIAL_RC − ST_INITIAL_RC, $\lambda_4 - \lambda_7$[e]	NA	NA	−0.10	−0.82
LT_SUB_NONPROX_RC−ST_SUB_NONPROX_RC, $\lambda_5 - \lambda_8$[e]	NA	NA	−0.37	−2.43**
LT_SUB_PROX_RC − ST_SUB_PROX_RC, $\lambda_6 - \lambda_9$[e]	NA	NA	−0.79	−5.50**

Panel C: Differences by Prior Restructuring Charge Proximity

Coefficient Difference	*n*[b]	*Median*[c]	*Weighted Mean Coefficient*	*Z-Statistic*[d]
LT_INITIAL_RC − LT_SUB_NONPROX_RC, $\lambda_4 - \lambda_5$[e]	NA	NA	−0.15	−0.74
LT_INITIAL_RC − LT_SUB_PROX_RC, $\lambda_4 - \lambda_6$[e]	NA	NA	−0.31	−1.86
LT_SUB_NONPROX_RC − LT_SUB_PROX_RC, $\lambda_5 - \lambda_6$[e]	NA	NA	−0.16	−0.53
ST_INITIAL_RC − ST_SUB_NONPROX_RC, $\lambda_7 - \lambda_8$[e]	NA	NA	−0.25	−1.29
ST_INITIAL_RC − ST_SUB_PROX_RC, $\lambda_7 - \lambda_9$[e]	NA	NA	−0.83	−4.08**
ST_SUB_NONPROX_RC − ST_SUB_PROX_RC, $\lambda_8 - \lambda_9$[e]	NA	NA	−0.58	−3.17**

---

*, ** Significant at 0.05 and 0.01, respectively.

[a] See Table 6 for a complete specification of Equation (5).

[b] n is the number of firm-specific parameters used to compute differences.

[c] The significance levels indicated in the Median column are based on the Wilcoxon signed rank test.

[d] Appendix A describes the weighted mean and Z-statistic calculations.

[e] We evaluate differences by CEO tenure and temporal proximity of the charge using the two-sample form of the z-statistic described in Appendix A. Sample sizes for each two-sample comparisons correspond to numbers of coefficient observations reported in Table 5. Medians are not reported for this analysis.

prior charges has a greater effect on CEO cash compensation when the CEO is short-tenured than when he is long-tenured.

Sensitivity Analyses

Prior research has found that CEO compensation is associated with the firm's stock returns as well as with the firm's earnings. We reestimate all models including three different annual stock return measures (raw, abnormal from market model using value-weighted indices, and abnormal from market model using equally weighted indices) as additional control variables. The results of these tests (not reported) indicate that CEO compensation increases with the firm's stock return, but our primary findings are robust to controlling for returns.

In addition to the firm-specific regressions reported in the tables, we also estimate all regressions using a pooled cross-sectional time-series design. To control for firm-specific and year-specific differences, we include dummy variables for each firm and year. The only difference between the pooled cross-sectional design results and the firm-specific design results is that the coefficient magnitudes in the pooled design are substantially smaller than those in the firm-specific design. However, the relations among the variables remains consistent with the firm-specific results reported in the tables (i.e., compensation committees, on average, partially shield CEO cash compensation from the effect of restructuring charges; the degree of shielding depends on whether the restructuring is an initial, temporally non-proximate subsequent, or temporally proximate subsequent restructuring charge; and the longer the executive's tenure, the more compensation committees shield CEO compensation from the effect of restructuring charges).

Prior research suggests that CEO cash compensation increases with earnings persistence (Baber et al. 1998, 1999). Because restructuring charges are generally nonrecurring, a control for earnings persistence may be particularly critical in restructuring years. To control for the potential effect of earnings persistence on the relation between CEO cash compensation and restructuring charges, we estimate a first-difference specification of Equation (2) including a firm-specific control for earnings persistence.[20] We estimate this specification using both a firm-specific and a pooled cross-sectional time-series design. Consistent with our basic analysis reported in Table 3, the results of these tests suggest that compensation committees partially shield CEO compensation from the earnings effect of restructuring charges, on average.

V. CONCLUSION

We reexamine the degree to which compensation committees appear to intervene to protect CEO compensation from the adverse effect of restructuring charges. Our results confirm Dechow et al.'s (1994) (DHS's) conclusions that compensation committees intervene to modify the income number used to determine CEO compensation, and that the extent of this intervention depends on the frequency of reported restructurings and the length of the CEO's tenure. However, in contrast to prior evidence that compensation committees fully shield executive cash compensation from the earnings effect of nonrecurring losses, including restructuring charges (DHS; Gaver and Gaver 1998), we find that after controlling for the temporal trend in inflation-adjusted CEO compensation, compensation committees, on average, only partially shield CEO compensation from the adverse effect of restructuring charges.[21] Thus, our results suggest that prior evidence that compensation committees completely shield CEO compensation from the effects of restructuring charges is likely a spurious result of a correlated omitted variable—the temporal increase in inflation-adjusted CEO cash compensation. This highlights the importance of controlling for the temporal trend in inflation-adjusted CEO compensation in firm-specific time-series compensation models.

---

[20] Following Ali and Zarowin (1992) and Cheng et al. (1996), we rank firms each year by their E_{it}/P_{it} ratio, where E_{it} is reported earnings per share for year t and P_{it} is the end-of-year stock price. We assign all firm-years with negative values of E_{it}/P_{it} a ranking of 1, and group the remaining firm-years into nine approximately equal portfolios and assign them rankings of 2 through 10. The earnings persistence variable equals 1 if the firm-year's E/P ratio falls in portfolios 3 to 8, 0 otherwise. This ranking procedure assumes the extreme portfolios (top two and bottom two) have more transitory items in earnings (i.e., earnings are less persistent) than the middle portfolios.

We also investigate how the context of the restructuring affects the extent of shielding. Our results support DHS's conclusion that, in general, compensation committees provide less shielding in firms reporting multiple restructuring charges, and more shielding in firms with long-tenured vs. short-tenured CEOs. Further, we find differential shielding for three specific types of restructuring charges—initial, nonproximate subsequent, and proximate subsequent charges—and that these various kinds of restructuring charges interact with CEO tenure. Specifically, our results suggest that compensation committees: (1) completely shield initial restructuring charges and subsequent restructuring charges not preceded by a prior restructuring within two years if the CEO has long tenure, (2) provide no shielding of subsequent restructuring charges that follow a prior charge within two years if the CEO has short tenure, and (3) partially shield the other categories of restructuring charges. To our knowledge, this is the first evidence that, depending on the context, compensation committees let some restructuring charges flow through, thereby reducing the earnings number used to compute CEO compensation.[22] In sum, this study demonstrates that contextual factors appear to play a substantive role in compensation committee decisions whether to shield (and if so how much to shield) executive compensation from the adverse effects of restructuring charges.

Our study has several limitations. First, we examine only CEO cash compensation. Recent research suggests that other forms of executive compensation, such as stock options, represent an increasing portion of total executive compensation in recent years (Murphy 1999). Furthermore, since firms typically restructure to improve *future* performance, total compensation may provide a more appropriate measure of whether compensation committees award CEOs additional incentive compensation for undertaking restructurings. We leave this question for future research. Second, the evidence that intervention is contextual is a necessary but not sufficient condition for demonstrating that compensation committees penalize CEOs for opportunistic restructuring charges. The study's evidence is indirect in at least two respects. First, we do not study compensation committees' deliberations; instead the analysis uses observable outputs of that process to draw inferences about the committees' decision rules. Second, we do not identify "opportunistic" restructurings. Perhaps future research could investigate more directly whether compensation committees are less apt to shield restructurings (or other actions) that are not value-enhancing, or that appear to be undertaken primarily to manage earnings. Finally, our sample includes only *Fortune* 500 firms. The degree of shielding we document may not generalize to smaller firms not in the *Fortune* 500. Despite these limitations, our results dispel prior evidence that compensation committees unconditionally and fully shield executive compensation from the adverse effect of restructuring charges on earnings.

---

[21] In fact, the DHS results imply that CEOs are awarded bonus pay for undertaking restructurings.

[22] One possible exception is Gaver and Gaver's (1998, 250) evidence that compensation committees partially shield executive compensation from the earnings effect of "income-decreasing special items." However, special item losses reported by Compustat include a variety of transactions, not just restructuring charges.

Appendix A
Derivation and Distribution of the Weighted Mean Statistic

We provide coefficient estimates consistent with the framework underlying the Z-statistics employed both here and in DHS. For a given independent variable, this estimate is a weighted mean of the underlying firm-specific coefficient estimates. The weighting follows directly from the assumptions underlying the Z-statistic. Specifically, the regression coefficient estimate, b_j, of each firm j is distributed as:

$$t(b, s_{bj}^2)$$

where:

$t(.)$ = the student's t distribution;
b = the population mean for the regression parameter; and
s_{bj}^2 = the variance of the firm-specific coefficient estimate.

Multiplying b_j by w_{bj}, where w_{bj} equals $s_{bj}\sqrt{(k_j/(k_j/(k/zj - 2)))}$ and k_j is the number of yearly observations in firm j's compensation regression, yields a value (asymptotically) distributed as:

$$\text{Normal } (w_{bj}\, b, 1).$$

The sum of these multiples across the sample is then distributed:

$$\text{Normal}\left(\sum_{j=1}^{N} w_{bj} b, N\right)$$

Dividing the sum of the $w_{bj}b_j$ multiples by $\sqrt{N}$ yields the conventional Z-statistic employed in DHS. Dividing this sum by $\Sigma_{j=1}^{N} w_{bj}$ yields the corresponding population weighted mean estimator, distributed as:

$$\text{Normal}\left(b, N\Big/\left(\sum_{j=1}^{N} w_{bj}\right)^2\right).$$

We use this form of the matched-pair Z-statistic for the statistical analyses presented in Tables 3 through 6 and in Table 7 Panel A.

Moreover, the difference, D, between two such weighted means, assuming independence, is distributed as:

$$\text{Normal}\left(D, N_1\Big/\left(\sum_{j=1}^{N_1} w_{bj}\right)^2 + N_2\Big/\left(\sum_{i=1}^{N_2} w_{bi}\right)^2\right)$$

where D is the difference in the two population means (i.e., $b_1 - b_2$), N_1 is the number of firms in sample j, and N_2 is the number of firms in sample i. An associated two-sample Z-test divides the estimated difference between the two weighted mean estimates by the square root of:

TABLE A1
Replication of Dechow et al. (1994)
Results from 90 Firm-Specific Regressions of CEO Cash Compensation on Income and Restructuring Charges, 1982–1989[a]

Panel A: $COMP_t = \beta_0 + \beta_1 ADJ\_INC_t + \beta_2 R\_CHARGE_t$

	Predicted Sign	*Median*[b]	*Weighted Mean Coefficient*[c]
β_0		522.33	
β_1	+	0.54**	0.81**
β_2	?	−0.27*	−1.13**
$\beta_1 - \beta_2$	+	0.72**	1.94**
Adjusted R^2		0.30	
Number of Observations		20	

Panel B: $COMP_t = \beta_0 + \beta_1 ADJ\_INC_t + \beta_2 R\_CHARGE_{it} + \beta_3 TREND_t$

	Predicted Sign	*Median*[b]	*Weighted Mean Coefficient*[c]
β_0		133.03	
β_1	+	0.58**	0.84**
β_2	?	0.29**	0.36**
β_3	+	252.11**	262.83**
$\beta_1 - \beta_2$	+	0.23*	0.48**
Adjusted R^2		0.51	
Number of Observations		20	

---

*, ** Significant at 0.05 and 0.01, respectively.

[a] We adjust all dollar amounts to 1989 equivalent dollars based on the CPI.

[b] The significance levels indicated in the Median column are based on the Wilcoxon signed rank test.

[c] Appendix A describes the weighted mean coefficient and Z-statistic calculations.

COMP = CEO cash compensation (salary and bonus) in inflation-adjusted dollars;
R_CHARGE = the restructuring charge in year t, in inflation-adjusted dollars;
ADJ_INC = INCOME − R_CHARGE, where INCOME equals earnings before tax, extraordinary items, and the results of discontinued operations in year t, in inflation-adjusted dollars; and
TREND = inflation-adjusted CEO cash compensation growth index.

$$N_1 \Big/ \left(\sum_{j=1}^{N_1} w_{bj}\right)^2 + N_2 \Big/ \left(\sum_{i=1}^{N_2} w_{bi}\right)^2.$$

We use this form of the Z-statistic for the empirical analyses reported in Table 7, Panels B and C.

Appendix B

Replication of Dechow et al. (1994)

We replicate the Dechow et al. (1994) (DHS) sample as completely as possible based on a listing of the firms comprising their sample that Professor Huson graciously provided us. We were able to locate cash compensation and restructuring charge data for all but one of the 91 firms in the DHS sample. Consistent with DHS, we estimate the compensation-earnings relation on a firm-by-firm basis using all compensation and earnings data available for the period from 1970 through 1989. Hence, there are a maximum of 20 time-series observations for each firm. Consistent with DHS, we require that the firm have a minimum of ten years of compensation data to remain in the test sample.

We report the replication analysis in Table A1. Panel A reports our results for the primary DHS time-series regression coefficient means and medians. Consistent with DHS, the average R_CHARGE coefficient is significantly negative and the average coefficient on ADJ_INC is significantly greater than the coefficient on R_CHARGE. Panel B reflects the effect of adding the compensation-trend variable to the basic DHS model. In this analysis, the average R_CHARGE coefficient is positive and significant (0.01 level). Consistent with our Table 3 analysis, we also find that the coefficient on ADJ_INC is significantly greater than the coefficient on R_CHARGE (0.01 level). Hence, we find that after controlling for the over-time increase in inflation-adjusted CEO cash compensation, on average, compensation committees partially shield executive compensation from the effect of restructuring charges.

REFERENCES

Ali, A., and P. Zarowin. 1992. The role of earnings levels in annual earnings-returns studies. *Journal of Accounting Research* 30 (Autumn): 286–296.

Atiase, R., D. Platt, and S. Tse. 2001. Corporate operational restructuring, financial and operating performance, and financial reporting. Working paper, The University of Texas at Austin.

Baber, W., S. Kang, and K. Kumar. 1998. Accounting earnings and executive compensation: The role of earnings persistence. *Journal of Accounting and Economics* 25 (May): 169–193.

———, ———, and ———. 1999. The explanatory power of earnings levels vs. earnings changes in the context of executive compensation. *The Accounting Review* 74 (October): 459–472.

Bliss, R. J., and R. T. Rosen. 2001. CEO compensation and bank mergers. *Journal of Financial Economics* 61 (July): 107–138.

Brickley, J. A., and L. D. Van Drunen. 1990. Internal corporate restructurings: An empirical analysis. *Journal of Accounting and Economics* 12 (January): 251–280.

Byrne, J. A. 1991. The flap over executive pay. *Business Week* (May 6): 90–110.

Carter, M. E. 2000. Operating performance following corporate restructurings. Working paper, Columbia University.

Cheng, G. S. A., C-S. Liu, and T. Schaefer. 1996. Earnings permanence and the incremental information content of cash flows from operations. *Journal of Accounting Research* 34 (Spring): 173–181.

Dechow, P., M. Huson, and R. Sloan. 1994. The effect of restructuring charges on executives' cash compensation. *The Accounting Review* 69 (January): 138–156.

Elliott, J., and W. Shaw. 1988. Write-offs as accounting procedures to manage perceptions. *Journal of Accounting Research* 26 (Supplement): 91–119.

Gaver, J., and K. Gaver 1998. The relation between nonrecurring accounting transactions and CEO cash compensation. *The Accounting Review* 73 (April): 235–254.

Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, New York, NY, September 28.

Lopez, T., P. Regier, and L. Holder-Webb. 2002. Do restructurings improve operating performance? Working paper, Georgia State University.

Moehrle, S. R. 2002. Do firms use restructuring charge reversals to meet earnings targets? *The Accounting Review* 77 (April): 397–413.

Muckian, M. 1993. CEO roundtable: CEO compensation continues to climb, survey says. *Credit Union Magazine* (July): 10.

Murphy, K. 1999. Executive compensation. In *Handbook of Labor Economics*, Vol. 3, edited by O. Ashenfelter, and D. Card. New York, NY: North Holland.

Pourciau, S. 1993. Earnings management and nonroutine executive changes. *Journal of Accounting and Economics* 16 (Jan-April-July): 317–336.

Pratt, N. 1996. CEOs reap unprecedented riches while employees' pay stagnates. *Journal of Compensation and Benefits Review* (September): 20–25.

Romani, P. 1997. CEO compensation: How much is enough? *Supervision* 58 (September): 3–6.

Smart, T. 1999. Pay gap between workers and bosses explodes. *The Washington Post* (August 30): A6.

Strong, J. and J. Meyer. 1987. Asset writedowns: Managerial incentives and security returns. *Journal of Finance* 42 (July): 643–663.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 193–225

Analyst Forecast Revisions and Market Price Discovery

Cristi A. Gleason
The University of Arizona
Charles M. C. Lee
Cornell University

ABSTRACT: We document several factors that help explain cross-sectional variations in the post-revision price drift associated with analyst forecast revisions. First, the market does not make a sufficient distinction between revisions that provide new information ("high-innovation" revisions) and revisions that merely move toward the consensus ("low-innovation" revisions). Second, the price adjustment process is faster and more complete for "celebrity" analysts (*Institutional Investor All-Stars*) than for more obscure yet highly accurate analysts (*Wall Street Journal Earnings-Estimators*). Third, controlling for other factors, the price adjustment process is faster and more complete for firms with greater analyst coverage. Finally, a substantial portion of the delayed price adjustment occurs around subsequent earnings-announcement and forecast-revision dates. Collectively, these findings show that more subtle aspects of an earnings revision signal can hinder the efficacy of market price discovery, particularly in firms with relatively low analyst coverage, and that subsequent earnings-related news events serve as catalysts in the price discovery process.

Keywords: *analysts; forecasts; revisions; market efficiency; earnings quality.*

Data Availability: *Data are available from sources identified in the paper.*

I. INTRODUCTION

This study examines the information content and market price discovery process associated with individual analysts' earnings forecast revisions. Our primary goal is to better understand the characteristics of forecast revisions that affect their immediate

---

We thank Ashiq Ali, Larry Brown, John Elliott, Laurie Krigman, Lillian Mills, Mark Nelson, Praveen Sinha, Bhaskaran Swaminathan, two anonymous reviewers, workshop participants at The University of Arizona, Cornell University, and the New York University Stern School, and especially Linda Bamber, for helpful comments. We also gratefully acknowledge the contribution of Thomson Financial Services Inc. for providing earnings per share forecast data, available through the Institutional Brokers Estimate System (I/B/E/S). These data have been provided as part of a broad academic program to encourage earnings expectation research.

Submitted October 2000
Accepted August 2002

market impact, as well as the magnitude of the subsequent price drift over the next year. We are particularly interested in how these characteristics affect the proportion of the total price response that occurs immediately, relative to the price drift that occurs in the following months.

Individual analysts' forecast revisions play an important role in the dissemination of information about corporate earnings. Unlike quarterly earnings announcements, individual analyst forecast revisions take place throughout the quarter. Because of their frequency and timeliness, these revisions have become a vital source of information for many users of corporate financial reports. Web-based technology now delivers individual forecast revisions to investors in real time.[1] As access to this information increases, the need to better understand the information content of each individual revision also increases.

Our study is motivated by the ongoing debate on the efficiency with which the market assimilates earnings news (e.g., Kothari 2001; Lee 2001). The delayed market response to earnings news is one of the most perplexing anomalies to emerge from accounting-based capital market research over the past 20 years.[2] Although this empirical regularity is well documented, surprisingly little is known about qualitative features of an earnings signal that can either hinder or aid the market in assimilating this information.

In this study, we identify and test several factors that we expect will affect the speed and efficacy of the price discovery process. Our investigation focuses on attributes of the forecast revision itself, as well as on characteristics of the firm for which the forecast is being made. Specifically, we hypothesize that, after controlling for the magnitude and direction of the revision surprise and for other firm characteristics, the magnitude of the post-revision price drift is related to (1) the extent to which a revision is innovative relative to existing forecasts (i.e., whether an analyst brings new information or is simply herding toward the consensus);[3] (2) the analyst's ability and reputation (i.e., whether he or she is a superior forecaster, and whether his/her abilities are likely to be well known to the market); and (3) analyst coverage (i.e., whether many analysts cover a firm).

Our results show that all three factors are related to the subsequent price discovery process. First, we find that the post-revision price drift is more pronounced for high-innovation revisions. Second, we find that forecasts made by celebrity analysts (*Institutional Investor [II] All-Stars*) elicit a stronger immediate price response, but a less pronounced subsequent price drift. In contrast, the subsequent price drift is more pronounced for revisions made by superior analysts who are more likely to be employed by smaller firms (*Wall Street Journal [WSJ] Earnings-Estimators*). Third, controlling for other factors, we show that the post-revision price drift is more pronounced in firms with lower analyst coverage.

In addition, we show that a significant portion of the delayed price reaction occurs around the next *four* earnings announcements. Interestingly, we find an even greater portion of the delayed reaction occurs in short windows around the next *six* forecast revisions. This evidence suggests that the market is adjusting slowly to the information contained in individual forecast revisions, and that subsequent revisions act as catalysts in the price discovery process.

---

[1] All the primary service providers (e.g., I/B/E/S, First Call, and Zacks) now offer a menu of web-based products that deliver real-time earnings estimate revisions. The Appendix provides a sample screen capture of First Call information delivered through FactSet, a high-end investment service provider.

[2] Ball and Brown (1968), Foster et al. (1984), Freeman and Tse (1989), and Bernard and Thomas (1989, 1990) document the post-earnings-announcement price drift; Givoly and Lakonishok (1980) and Stickel (1991) document a similar drift after revisions in analysts' earnings forecasts. These studies form the basis of earnings-based momentum strategies that many investment firms use.

[3] As explained later, we distinguish between high-innovation and low-innovation revisions using two benchmarks: an analyst's own prior forecast and the prior day's consensus forecast. Specifically, we define ***high-innovation*** forecasts as those that are either higher or lower than *both* benchmarks. Conversely, *low-innovation* revisions are those that are between the analyst's own prior forecast and the prior consensus estimate. As we define them, high-innovation revisions bring unambiguous information to market while low-innovation revisions could simply reflect herding toward the consensus.

Our findings point to several imperfections in the price discovery process for earnings forecast revisions. First, the market does not make a sufficient distinction between analysts who are unambiguously providing new information and those who are simply herding toward the consensus. After we control for the magnitude of the forecast revision, high-innovation revisions are associated with significantly larger price drifts than low-innovation revisions. This result suggests that the level of innovation is an aspect of the revision signal that the market does not fully incorporate at the time of the announcement. Second, the market pays more attention to analysts who have acquired celebrity status, but is more likely to underappreciate revisions made by more obscure analysts with comparable forecasting abilities. This result suggests that an analyst's reputation, and not merely his forecasting ability, affects market price discovery. Finally, the price drift is lower for firms followed by more analysts, suggesting that coverage by multiple analysts helps to facilitate the price discovery process.

Our study contributes to several streams of recent research. First, it extends the growing literature on the role of analysts as information intermediaries. Prior studies find evidence that the level of financial analyst coverage affects the efficiency with which the market processes information (Brennan et al. 1993; Walther 1997; Bhattacharya 2001; Elgers et al. 2001). We show that increased analyst coverage generally results in a lower price drift subsequent to analyst forecast revisions, suggesting that in the case of individual forecast revisions, additional coverage helps to facilitate price discovery.

Second, our results highlight the usefulness of analyst information in investment decisions. Elgers et al. (2001) report that a simple trading strategy based on the ratio of price to one-year-ahead earnings forecasts (P/FY1) yields significant positive abnormal returns. Other studies (e.g., Womack 1996; Jegadeesh et al. 2001) find that analyst stock-recommendation revisions predict returns over the next 3 to 12 months. Our findings show that one can enhance the forecast revision investment strategy suggested by Stickel (1991) by considering the level of innovation of the revision signal, the amount of analyst coverage, and, to a lesser extent, the ranking of the analyst.

Finally, our results shed light on the price discovery mechanism and help define the boundaries of market efficiency. Much of the regulatory debate in accounting standard setting revolves around the issue of "transparency"—that is, whether investors can readily decipher more subtle aspects of financial reports. Our findings suggest that market participants do not immediately assimilate certain potentially informative, qualitative aspects of an earnings signal. By inference, these findings raise questions as to the market's ability to quickly and accurately incorporate other, more subtle, aspects of the quality of reported earnings. Collectively our results imply that the price discovery process, at least as it pertains to analyst forecast revisions, is complex and protracted, and deserving of more scholarly attention.

II. RELATED LITERATURE AND HYPOTHESES DEVELOPMENT

Related Literature

Two main stylized facts emerge from prior research on analyst forecast revisions. First, these forecast revisions are price informative, in that they are associated with a significant immediate market response at the time of their release. Second, the immediate price response is incomplete, in that prices, on average, continue to drift in the same direction for at least three to nine months after the price revision.

Evidence that analyst forecasts are price-informative dates back to the 1970s. Early research by Griffin (1976), Givoly and Lakonishok (1979, 1980), and Imhoff and Lobo (1984) documents the market response to analyst forecast revisions. In fact, Elton et al. (1981) show that foreknowledge of analyst revisions is more value-relevant than foreknowledge of the reported earnings themselves. Collectively, these findings show that individual analyst forecast revisions convey new information to the market.

Subsequent research shows that the short-window price reaction around forecast revisions is incomplete. Using a small sample, Givoly and Lakonishok (1980) first reported a post-announcement drift after earnings forecast revisions. Stickel (1991) demonstrated that firms whose consensus forecast has been recently revised upward tend to earn higher abnormal returns over the next 3 to 12 months than firms whose consensus forecast has been recently revised downward. Finally, Chan et al. (1996) confirm Stickel's (1991) finding and show that it is part of a general class of "momentum" strategies, in which the market response to recently released information is incomplete.[4]

Although the post-revision price drift is well documented, surprisingly little is known about factors that either exacerbate or mitigate this empirical regularity. Our study focuses on factors that help explain cross-sectional variation in the delayed price reaction. Unlike most prior studies, which focus on changes in the *consensus* analyst forecast, we examine the market response to *individual* analyst revisions. This approach allows us to evaluate qualitative attributes of the forecast revision itself, as well as elements in the firm's information environment that could potentially affect the information-assimilation process.

Hypothesis Development

Figure 1 provides an overview of the key conceptual constructs that we examine in this study, and of the empirical proxies we employed for each construct. As this figure illustrates, our investigation focuses on four sources of potential variation in the delayed price response. Our main interest lies in uncovering *qualitative* features of the forecast revision signal and elements in the firm's information environment that affect the efficacy of the information-assimilation process. However, we recognize the need to control for the *quantity* of the forecast revision (i.e., its direction and magnitude) as well as for various *firm characteristics* that can affect expected returns. We discuss each of our four key conceptual constructs in more detail below.

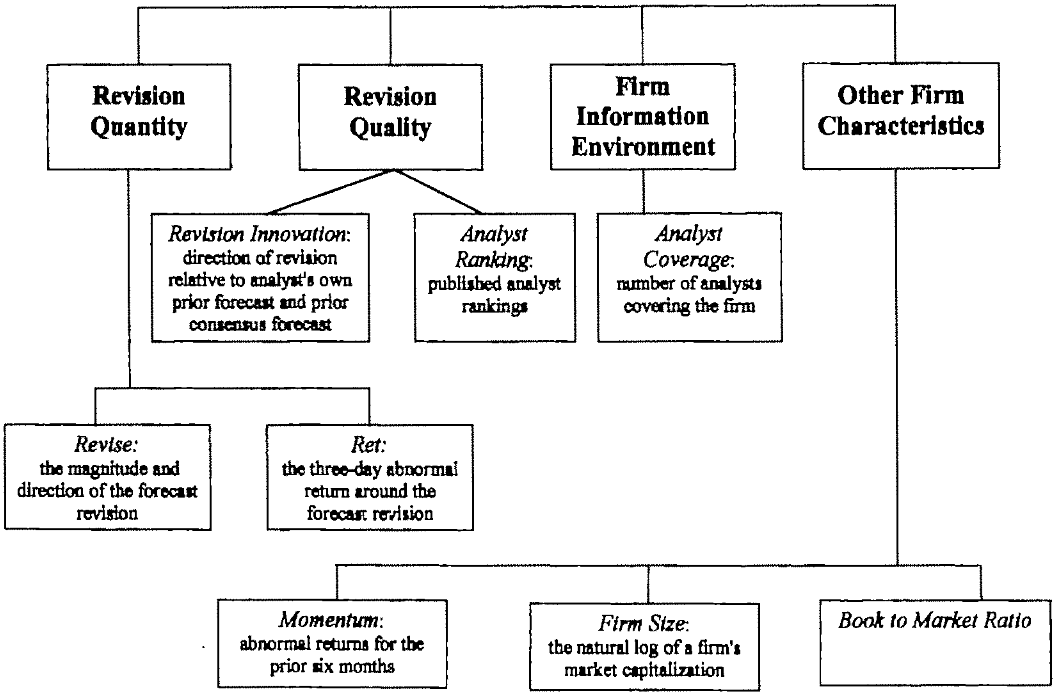
Revision Quantity

First, we examine the conjecture that the magnitude of the post-revision drift is a function of the size and direction of the revision signal itself. Prior studies on the post-earnings-announcement drift document a positive relation between the *quantity* of the unexpected earnings news (i.e., the size and direction of the earnings surprise) and the subsequent price drift (e.g., Bernard and Thomas 1990). Although the effect these studies document pertains to actual earnings surprises rather than analyst forecast revisions, we hypothesize that an earnings revision of greater magnitude will also trigger a greater subsequent price drift.

---

[4] Others who have documented similar price drifts after different news events include Bernard and Thomas (1989, 1990) and Freeman and Tse (1989), who find price drifts after earnings surprises, and Womack (1996), who documents a price drift after analysts' buy/sell recommendations. In each case, the evidence points to an immediate short-window price response to the information release, followed by a generally sluggish adjustment in the same direction.

FIGURE 1
Factors Contributing to the Delayed Price Reaction to Analyst Forecast Revisions


Revision Quantity
Revision Quality
Firm Information Environment
Other Firm Characteristics
Revision Innovation: direction of revision relative to analyst's own prior forecast and prior consensus forecast
Analyst Ranking: published analyst rankings
Analyst Coverage: number of analysts covering the firm
Revise: the magnitude and direction of the forecast revision
Ret: the three-day abnormal return around the forecast revision
Momentum: abnormal returns for the prior six months
Firm Size: the natural log of a firm's market capitalization
Book to Market Ratio

We investigate this hypothesis using two empirical proxies for the quantity of the forecast revision signal. The first measure, Revise, is the magnitude of the revision expressed as a percentage of the firm's closing price on day −1 relative to the revision. Our choice of this variable is based on prior studies (e.g., Stickel 1991; Imhoff and Lobo 1984) that show an analyst's own prior forecast is a better benchmark than the firm's consensus forecast for measuring the amount of surprise in the individual forecast. Our second measure of revision quantity, Ret, is the firm's size-adjusted return over the three-day window surrounding the revision. The immediate market response to a forecast revision is arguably a less noisy measure of the degree of surprise contained in the revision.

Revision Quality

We also investigate two qualitative attributes of the forecast revision signal, to ascertain whether nonquantitative characteristics of the forecast revision provide value-relevant information that is not fully reflected in the three-day returns over the event window. For example, Sloan (1996) documents a delayed market response to value-relevant information concerning earnings quality contained in total accruals. In the same spirit, we examine more subtle aspects of the forecast revision signal that might help explain cross-sectional differences in the delayed market reaction.

Revision Innovation

The first qualitative attribute we examine is the level of innovation in the forecast revision. Analyst forecasts are typically reported along with the analyst's prior forecast and the most recent consensus forecast. Therefore, both the analyst's own prior forecast and the prior consensus forecast are potentially important benchmarks in assessing the information content of the revision. We use both benchmarks in distinguishing between *high-innovation* and *low-innovation* revisions.

Figure 2 illustrates the difference between high- and low-innovation revisions. In this figure, we label forecasts that are higher than both the analyst's own prior forecast and the prior consensus as *high-innovation good-news* revisions (denoted G). Similarly, we label forecasts that are lower than both the analyst's own prior forecast and the prior consensus as *high-innovation bad-news* revisions (denoted B). The direction of a high-innovation revision relative to both benchmarks suggests unambiguously that the analyst is bringing new information to the market. In contrast, we label forecasts that are between the analyst's own prior forecast and the prior consensus *low-innovation* revisions. For low-innovation revisions, whether the revision indicates good news (denoted L) or bad news (denoted L′) depends on whether we compare the revision to the analyst's own prior forecast or the prior consensus. When the new forecast is exactly the same as the prior consensus, we treat it as a low-innovation revision.

An earlier finding by Stickel (1990) suggests that low-innovation forecasts likely contain less information than high-innovation forecasts. Specifically, Stickel (1990) shows that changes in the consensus are useful in predicting the next revision of any given analyst. This result implies that low-innovation revisions are less informative (regardless of magnitude), because they simply reflect information revealed by the prior revisions of other analysts. We therefore hypothesize that a high-innovation revision will trigger a greater immediate price impact than a low-innovation revision.

We are more interested, however, in the effect of revision innovation on the market price discovery process. The level of revision innovation is a qualitative, rather than quantitative, characteristic of the forecast revision. Consequently it is more subtle in nature. We conjecture that the delayed price reaction is a function of the level of innovation in the revision signal. One could argue that analysts who revise their forecasts *toward* the consensus forecast are "herding" rather than bringing new information to market. To the extent that the market does not sufficiently distinguish between those analysts who herd vs. those who bring new information to market, we expect high-innovation revisions to be associated with a greater subsequent price drift.[5]

Analyst Ranking

The analyst's ranking is another characteristic of forecast revisions that affects both the information content of the signal and the efficiency of the price discovery process. Early research found no systematic differences among analysts in terms of the accuracy of their forecasts (e.g., Brown and Rozeff 1980; O'Brien 1990; Butler and Lang 1991). However, Stickel (1992) uses more comprehensive controls for the age of the forecasts and shows that analysts on the *II All-Star* team provide more accurate forecasts. Using a large sample that spans many industries, Sinha et al. (1997) also show that certain analysts consistently outperform others in terms of the accuracy of their forecasts.

Subsequent studies have focused on the economic determinants of individual analyst forecast accuracy, such as analyst experience (Mikhail et al. 1997); analyst abilities, resources, and portfolio complexity (Clement 1999); analyst aptitude and brokerage house

---

[5] We further decomposed high-innovation revisions into "diverging" signals (those that move away from the consensus) and "converging" signals (those that move toward, and beyond, the consensus). However, this additional partition had little effect on our results, and we do not separately report the findings.

FIGURE 2
Classification of Individual Forecast Revisions

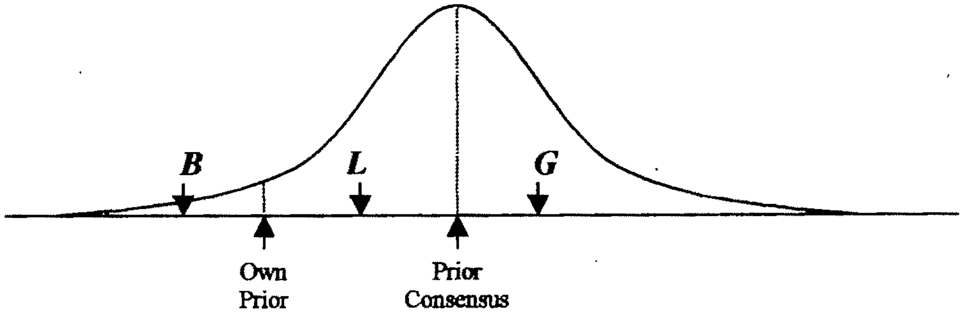


Figure 2a: When an analyst's own prior forecast is less than the current consensus

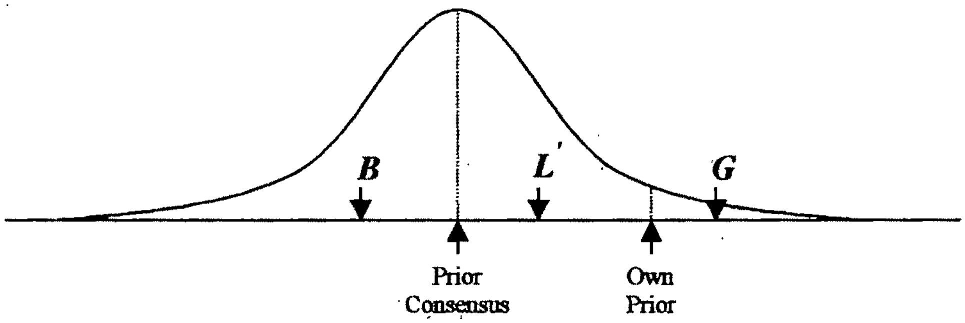


Figure 2b: When an analyst's own prior forecast is greater than the current consensus

Four Possible Classifications:

- *B – High-innovation signal, bad news.*
- *L – Low-innovation signal, good news.*
- *L' – Low-innovation signal, bad news.*
- *G – High-innovation signal, good news.*

characteristics (Jacob et al. 1999); and age of the forecast (Brown and Mohammad 2001). Brown (2001) demonstrates that past accuracy is as or more useful in predicting future forecast accuracy as the economic determinants identified in the literature just mentioned. All these studies focus on the determinants and persistence of individual analyst accuracy, rather than on the market response to individual forecasts.[6]

Our research focuses on analyst attributes that could affect both the reliability and the salience of a forecast revision to market participants. Specifically, we use two widely publicized sources to identify award-winning analysts: (1) the *Institutional Investor's All-American Research Team* (*II All-Stars*), and (2) the *Wall Street Journal's Survey of Award Winning Analysts*. Specifically, we focus on those analysts the *WSJ* identified as superior in terms of their ability to forecast earnings accurately (*WSJ Earnings-Estimators*). Stickel (1992) used the former in his study, and the latter is a more recent survey that has, to the best of our knowledge, not been used in prior research.

Stickel (1992) shows that the earnings forecasts of *II All-Stars* are, on average, more accurate than those of unranked analysts. However, the *II All-Stars* are selected by a survey of major institutional investors. According to *II*, selection is based on the economic insight in their written reports, the level of access and the overall service they provide to clients, and the quality of their stock recommendations, as well as the accuracy of their earnings forecasts. In short, these are celebrity analysts whose overall reputations might affect the way market participants perceive their revisions.

The *WSJ Earnings-Estimators* offer an interesting contrast to the *II All-Stars*. Like *II All-Stars*, these analysts are superior earnings forecasters. Whereas *II* selects its *All-Stars* based on a survey of buy-side clients, the *WSJ* bases its selection process strictly on forecast accuracy. *All-Stars* are generally from major investment houses, while the *WSJ* analysts often come from smaller and lesser-known firms. *All-Star* rankings are also sticky, in that once an analyst acquires *All-Stars* status, he or she is likely to retain this distinction in subsequent years. The turnover is much higher among *WSJ* rankings, which are based strictly on accuracy over the past 12 months. Therefore, although both sets of analysts have superior forecasting ability, the *WSJ* analysts typically do not enjoy the stature and overall reputation of the *II All-Stars*.

Prior studies show that the market differentiates these analysts to some extent in that superior analysts elicit stronger market responses in short windows immediately surrounding their forecast revisions (e.g., Stickel 1992; Park and Stice 2000). However, we are unaware of any study that examines the relation between analyst ranking and the delayed market response to forecasts. Like revision innovation, the stature and reputation of the analyst are more subtle, qualitative aspects of a revision signal. To the extent that the celebrity status of the *All-Stars* helps the market to recognize the value of their forecasts, we expect these revisions to trigger smaller post-revision price drifts. Conversely, if the market does not fully appreciate the greater accuracy of *WSJ* analysts, their forecasts will trigger more pronounced price drifts.[7]

Firm Information Environment

The overall level of information available about the firm also influences the value of an individual analyst forecast. In particular, heavier analyst coverage is associated with faster and more complete price adjustment. For example, Brennan et al. (1993) report that stocks with greater analyst coverage react faster to market-wide common information compared to those with less analyst coverage. Hong et al. (2000) find that returns to momentum-based investment strategies are higher for firms with low levels of analyst coverage. Elgers et al. (2001) find that the price response to value-relevant information in analyst earnings forecasts is less complete for firms with lower levels of analyst coverage. In contrast to these prior studies, we focus on the price drift after analysts' forecast revisions.[8] If the findings in Elgers et al. (2001) apply to analysts' revisions, we would then expect a more delayed price response in low-coverage firms.

---

[6] A concurrent study by Clement and Tse (2003) examines the relation between factors that predict analyst accuracy and the market response to forecast revisions. Their study focuses on returns in the immediate event window, while our main interest is on factors that help explain the post-revision price drift.

[7] Our hypotheses are consistent with experimental evidence from the individual decision-making literature (e.g., Griffin and Tversky 1992; Bloomfield et al. 2000) showing that, when an investor has imperfect information about signal reliability, the salience of the signal will affect the individual's perception of its reliability.

Other Firm Characteristics

We also include three firm-level variables to control for cross-sectional differences in expected returns: price momentum (measured as the market-adjusted return over the prior six months), firm size (defined as the natural log of a firm's market capitalization measured as of the end of the last calendar year), and the book-to-market ratio (measured as of the end of the last calendar year). Earlier research shows that, on average, smaller firms with higher book-to-market ratios and more positive price momentum earn higher subsequent returns (e.g., Fama and French 1992; Jegadeesh and Titman 1993). We control for these variables to isolate the incremental ability of forecast revisions to predict cross-sectional returns.

III. SAMPLE SELECTION AND RESEARCH METHOD

Sample Selection Criteria

We obtained our analyst forecast data from the I/B/E/S Detail File, and our returns data from the Center for Research on Security Prices (CRSP) database. We use names from the I/B/E/S Analyst and Broker Name File to identify individual I/B/E/S analysts, and match these names to the award winners from *II* and *WSJ* publications.

Our sample period begins in October 1993, the first month that both *II* and *WSJ* provided a list of ranked analysts, and ends on December 31, 1998. We focus on forecasts of the one-year-ahead annual earnings (FY1). Our initial sample of 372,021 observations consists of FY1 forecast revisions with a prior forecast by the same analyst and CRSP return data for the firm.

Table 1 provides descriptive information on our sample firms. Panel A shows that, except for 1993 (for which we have only fourth-quarter data), our sample includes more than 4,000 firms per year. Excluding 1993, the average number of revisions per firm ranges from 15.0 (1996) to 18.5 (1998). However, the distribution is skewed, and the median number of revisions per firm is only 7 (1996) to 9 (1998).

Panel B of Table 1 reports the number of revisions in our sample, classified by firm size decile and stock exchange. We base threshold values for firm size deciles on beginning-of-year market capitalization for all New York Stock Exchange (NYSE) firms. Approximately two-thirds of the revisions are for NYSE-listed stocks. As expected, the majority of the forecast revisions in our sample are for larger firms, which attract greater analyst coverage. However, firm-level data show that the top six NYSE size deciles are evenly represented in our sample.

In subsequent tests, we compute size-adjusted returns for each firm by subtracting the mean buy-and-hold return of an equal-weighted portfolio of firms in the same NYSE size

---

[8] Elgers et al. (2001) investigate early-in-the-year price-to-earning-forecasts (a type of PE ratio), rather than changes in individual analyst forecast revisions.

TABLE 1
Descriptive Statistics for Sample Firms and the Number of Forecast Revisions by Year, Firm Size, and Stock Exchange

Panel A: Number of Firms and Revisions Per Year

	Number of Firms	Number of Revisions	Number of Revisions per Firm: Average	Median	Std. Dev.
1993 (Oct.–Dec.)	3,000	19,580	6.53	4	8.21
1994	4,024	63,147	15.69	8	21.43
1995	4,244	67,192	15.83	8	21.88
1996	4,756	71,083	14.95	7	21.48
1997	4,755	72,839	15.32	8	21.01
1998	4,224	78,180	18.51	9	26.01
Total	6,505	372,021	14.91	7	21.54

Panel B: Number of Revisions and Firms by Firm Size Decile and Stock Exchange

NYSE Size Decile[a]	Revisions: NASD	NYSE	Total	% of Sample	Firms: Number of Firms	% of Total Firms
10 (largest)	22,333	156,795	179,128	48.1	810	12.4
9	23,026	51,200	74,226	20.0	819	12.6
8	20,052	21,324	41,376	11.1	824	12.7
7	16,967	10,611	27,578	7.4	890	13.7
6	14,865	5,554	20,419	5.5	860	13.2
5	10,155	3,074	13,229	3.6	806	12.4
4	7,490	1,493	8,983	2.4	641	9.9
3	4,131	791	4,922	1.3	471	7.2
2	1,542	196	1,738	0.5	294	4.5
1 (smallest)	362	60	422	0.1	90	1.4
	120,923	251,098	372,021	100.0	6,505	100.0
	32.5%	67.5%				

---

[a] Threshold values for NYSE Size Deciles are based on beginning-of-year market capitalization for all NYSE firms.

decile over our holding period. If a stock is de-listed during the return accumulation period, then we assume the proceeds are reinvested to earn the average return of the matching size decile portfolio.

Computation of Explanatory Variables

As discussed, our tests call for a comparison of each forecast revision to an analyst's own prior forecast, as well as to the most recent prior consensus forecast. I/B/E/S reports a monthly consensus measure in the third week of every month, but does not update this measure on a daily basis. To facilitate a daily comparison, we construct a new consensus forecast each day by taking the average of all the individual forecasts available as of the end of the previous day. For those days on which I/B/E/S does compute a consensus, we checked our consensus forecast measure against the I/B/E/S measure and found the two variables to be closely matched.[9]

Our tests also require that we identify analysts ranked by *II* and *WSJ*. The Analyst and Broker Name File from I/B/E/S contains the last name and first initial of each analyst, as well as a unique analyst code and the code for the brokerage firm. In those few instances in which two or more analysts had the same last name and first initial, we used the industry and brokerage firm information to distinguish between the analysts. We then identified superior analysts based on the list of names in the award issue of the *II* or *WSJ* publications.

Institutional Investor publishes its analyst rankings annually in October, and identifies four award winners for each industry. We consider analysts ranked first, second, third, or runner-up in the annual *II* list as *All-Stars* for the 12 months following their inclusion in the survey. The *WSJ* began publishing its own independent ranking of sell-side analysts in September 1993. Each subsequent year, beginning with 1994, *WSJ* published this ranking in June. We classify analysts ranked first through fifth as *Earnings-Estimators* for the 12 months following their inclusion in the industrial rankings.[10]

We also examine the effect of analyst coverage, a firm-specific characteristic, on the price discovery process. We measure the number of analysts covering a firm as the maximum number of analysts with outstanding forecasts at any point during the year. Following Elgers et al. (2001), we assigned firms to the high-coverage category if the number of analysts following the firm is above the median for our sample in a given year, and to the low-coverage category if the number of analysts is below the median.

Econometric Issues

In most of our empirical tests, the unit of observation is an individual forecast revision. We believe that revision-level analysis is appropriate (and necessary) if we are to better understand the price discovery process associated with the release of each new forecast revision. However, to the extent that individual revisions do not occur randomly across firms and over time, the observations in a revision-level regression are not independent. Moreover, because individual revisions exhibit positive serial correlation (Stickel 1990), test statistics from an ordinary least squares (OLS) regression will be inflated.

We deal with this sample independence problem in several ways. First, in our revision-level regressions (Table 5), we compute the t-statistics using the Huber-White estimator described in Huber (1967), White (1980), and Diggle et al. (1994). This maximum-likelihood estimation procedure addresses cross-correlation problems arising from having multiple observations from the same firm. Specifically, in constructing the variance and covariance matrix, this procedure assumes (and estimates) a common component for all the forecast revisions that come from the same firm. Individual revisions from the same firm contribute collectively to the variance and covariance estimate for this firm-group. The estimated coefficients from this procedure are asymptotically consistent, and the Huber-White test statistics are an order of magnitude lower than test statistics from ordinary OLS regressions. This approach retains the power of using individual analyst-level data while ensuring that firm-level cross-correlations do not inflate our test statistics.

---

[9] We are interested in mimicking the consensus forecast that is broadly available to users on a daily basis, rather than in creating a superior consensus that includes only the most recent forecasts. Therefore, we do not attempt to eliminate "stale" forecasts. To the extent that stale forecasts introduce noise into the consensus forecast, they reduce the power of our tests.

[10] The *WSJ* distinguishes between analysts who excel at forecasting earnings and analysts who excel in stock picking. We focus on the top *Earnings-Estimators*. Analysts included in the September 1993 rankings are classified as superior earnings forecasters from October 1993 through June 1994.

Second, to alleviate concern over closely clustered revisions, we reestimated the analyses reported in all our tables using only revisions for firms that have no other revision occurring within a three-day window (days -1 to $+1$).[11] We have just over 150,000 such revisions. This procedure induces a "peek-ahead" bias because it requires that we know another revision is forthcoming when the first one was issued. Nevertheless, our main results are unchanged using this restricted sample.

Finally, we entirely eliminate the sample dependency problem arising from multiple revisions from the same firm by conducting a firm-level rather than revision-level analysis. To perform this analysis, we convert revision-level information into firm-level indicator variables, and regress future size-adjusted returns on these variables. We find that the firm-level analysis (reported later in Table 7) leads to similar inferences.

IV. EMPIRICAL RESULTS

Alternative Benchmarks for Measuring Forecast Revisions

Table 2 presents a comparison of mean size-adjusted returns for revisions classified using either the prior consensus forecast or an analyst's own prior forecast. This table reports the cumulative size-adjusted returns for the event window from day -1 to $+1$ (three days). It also shows the mean size-adjusted return in the period after the announcement window. We accumulate these returns over various holding periods, beginning on day -1 and ending on trading day $+64$ (three months), $+127$ (six months), $+190$ (nine months), and $+253$ (one year).[12]

As documented by prior studies (e.g., Givoly and Lakonishok 1980; Stickel 1991), Table 2 shows that most of the price response is delayed. On average, a firm that has a downward revision experiences a further downward price drift after the revision window of 3.4 to 4.4 percent per year. Upward revisions experience an upward price drift of 3.1 to 3.6 percent. In fact, in our sample only 18.5 (0.015/0.081) to 20.8 percent (0.021/0.101) of the total one-year price reaction occurs immediately in the three-day window surrounding the forecast revision.[13]

Table 2 confirms prior studies (e.g., Stickel 1991; Imhoff and Lobo 1984) showing that an analyst's own prior forecast is a better benchmark than the consensus forecast for measuring the amount of surprise in an individual forecast revision. Specifically, the last column of this table shows that mean size-adjusted returns are larger when we classify revisions relative to an analyst's own forecasts. This result holds for all return accumulation periods. Therefore, to secure a better measure of the revision signal, we compute the forecast revision (hereafter referred to as Revise) as the new forecast minus the analyst's own prior forecast, scaled by price.

Tabular Classifications

As discussed, our main interest centers on the empirical association between the delayed price response and revision and firm characteristics. As a first step, we report in a tabular format the size-adjusted returns for forecast revisions classified by characteristics of interest

---

[11] The mean (median) number of days between revisions is 12.75 (5) days. The length of time between forecasts is at least one day for 73 percent of the sample revisions.

[12] When comparing a revision to the prior consensus forecast, we exclude 1,322 observations with forecast revisions equal to the prior consensus forecast. Similarly, when comparing a revision to the analyst's own prior forecast, we exclude 1,923 observations with forecast revisions equal to the analyst's own prior forecast.

[13] We also examined returns over two-year holding periods. We do not report these results because almost all of the price drift occurs in the first year. Moreover, our two-year holding period results are sensitive to the inclusion or exclusion of 2000 (see Barber et al. 2003).

TABLE 2
Comparison of Mean Size-Adjusted Returns for Forecast Revisions Classified Relative to the Prior Consensus Forecast or the Analyst's Own Prior Forecast

	Revision Signal Relative to Prior Consensus Forecast			*Revision Signal Relative to Analyst's Own Prior Forecast*			
	Bad News	*Good News*	*Difference*	*Bad News*	*Good News*	*Difference*	*Difference in Differences*
Number of Forecasts	228,811	141,888		206,635	163,463		
Average Forecast Change[a]	−0.259	0.117		−0.185	0.093		
Size-Adjusted Returns[b]							
3 days	−0.010*	0.005*	0.015*	−0.013*	0.007*	0.021*	0.006*
3 months	−0.030*	0.018*	0.048*	−0.037*	0.021*	0.058*	0.009*
6 months	−0.037*	0.023*	0.059*	−0.045*	0.026*	0.072*	0.013*
9 months	−0.039*	0.028*	0.067*	−0.050*	0.034*	0.085*	0.018*
1 year	−0.044*	0.038*	0.081*	−0.057*	0.045*	0.101*	0.020*
Post-revision price drift	−0.034*	0.031*	0.065*	−0.044*	0.036*	0.080*	0.014*

---

* Indicates two-tailed t-test of difference from zero is significant at $\alpha < 0.001$.

[a] Forecast Change is the new forecast less either the prior consensus or the analyst's own prior forecast.

[b] Size-Adjusted Returns are the buy-and-hold return of the firm minus the buy-and-hold return for an equal-weighted portfolio of firms in the same NYSE size decile. Return accumulation periods are three days (day −1 to +1), three months (day −1 to +64), six months (day −1 to +127), nine months (day −1 to +190), and one year (day −1 to day +253), where days are trading days and day 0 is the day of the forecast revision. The post-revision price drift is measured from day +2 to day +253.

(Table 3). These results are only suggestive, because they do not control for the other cross-sectional factors that might affect the magnitude of the delayed price response. Nevertheless, they offer preliminary insights into the relative importance of our key variables in explaining cross-sectional variations in the delayed price response. These tabular results also allow us to observe the proportion of the total price response accumulated at various intervals after the revision.

Table 3 documents the mean size-adjusted returns for forecast revisions classified by the level of revision innovation, the analyst's ranking, and the extent of analyst coverage of the firm. Panel A presents results for high-innovation vs. low-innovation revisions for the full sample.[14] Panels B through D report the effect of analyst ranking, controlling for the level of revision innovation. Finally, Panels E and F report the effect of analyst coverage, controlling for the level of revision innovation. For parsimony, we do not report results of three-dimensional classifications, or the two-way interaction between analyst ranking and firm coverage.[15]

Three key results emerge from Table 3. First, the price adjustment process is more protracted for high-innovation revisions. Panel A shows that high-innovation revisions trigger a stronger immediate response: the mean three-day hedge return to high-innovation revisions is 2.3 percent, compared to 1.1 percent for low-innovation revisions (p-value < 0.0001).[16] The post-revision price drift is also larger for high-innovation revisions: the hedge return after the event window is 9.4 percent for high-innovation revisions, and only 3.3 percent for low-innovation revisions (p-value < 0.0001). More importantly, the *proportion* of the total one-year price response accounted for by the three-day event window is lower for high-innovation revisions: 20 percent vs. 24 percent for low-innovation revisions (p-value < 0.0001).

Second, this table provides limited evidence that analyst rankings can also affect price discovery. Comparing Panels B and C of Table 3, revisions by *All-Stars* trigger a greater three-day market response than those by unranked analysts. This result holds for both low- and high-innovation revisions (p-value < 0.0001 and p-value < 0.014, respectively). The *proportion* of the total one-year price response captured by the three-day event window is 23 to 40 percent for *All-Stars*, and only 19 to 21 percent for unranked analysts. However, the difference is statistically insignificant. Comparing Panels B and D, revisions by *Earnings-Estimators* also trigger a greater immediate market response than revisions issued by unranked analysts (p-values < 0.015). However, the *proportion* of the total one-year price response captured by the three-day event window for *Earnings-Estimators* is also not significantly different from those for unranked analysts.

Finally, Panels E and F of Table 3 show that analyst coverage is associated with cross-sectional variation in the delayed price response. The total post-revision price drift over the next 12 months is much larger for firms with low analyst coverage than for firms with high analyst coverage (p-value < 0.0001). Further tests (not reported) show this result is also robust across different analyst ranking categories. Moreover, the three-day returns capture

---

[14] We exclude the 1,923 observations with forecast revisions equal to the analyst's own prior forecast, which were also excluded from Table 2. We also exclude 6,535 observations where the analyst's own prior forecast equals the prior consensus forecast. Results are not sensitive to this restriction.

[15] As we show later in the multiple regression analysis, these additional cross-classifications do not yield significant incremental insights.

[16] We compute t-statistics for the hedge portfolio returns in Table 3 assuming equal weight for each forecast revision. Returns associated with downward revisions are multiplied by -1 to simulate short positions. To avoid outliers caused by small denominators, we compute Wilcoxon sign rank statistics when comparing the *proportion* of 12-month returns accumulated during the three-day announcement window. The reported p-values are based on two-tailed tests of either the t-statistics (for hedge returns) or Wilcoxon statistics (for the proportion tests).

TABLE 3
Comparison of Mean Size-Adjusted Returns for Forecast Revisions Classified by Revision Innovation, Analyst Ranking, and the Extent of Analyst Coverage

	Low-Innovation Revisions^a				*High-Innovation Revisions*^{a}			
	Bad News	*Good News*	*Difference*	*% of 1-year Return*	*Bad News*	*Good News*	*Difference*	*% of 1-year Return*
Panel A: Full Sample								
No. of Forecasts	32,142	54,032			170,961	106,428		
Average Forecast Change^b	−0.08	0.08			−0.19	0.10		
Size-Adjusted Returns^c								
3 days	−0.005*	0.006*	0.011*	24	−0.015*	0.008*	0.023*	20
3 months	−0.015*	0.005*	0.021*	45	−0.041*	0.028*	0.069*	58
6 months	−0.022*	0.006*	0.029*	62	−0.050*	0.036*	0.085*	72
9 months	−0.027*	0.014*	0.040*	88	−0.054*	0.044*	0.098*	83
1 year	−0.027*	0.019*	0.046*	100	−0.062*	0.056*	0.118*	100
Post-revision price drift	−0.021*	0.012*	0.033*		−0.048*	0.046*	0.094*	
Panel B: Non-Ranked Analysts^d								
No. of Forecasts	24,177	40,194			128,472	79,779		
Average Forecast Change^b	−0.08	0.07			−0.19	0.10		
Size-Adjusted Returns^c								
3 days	−0.005*	0.005*	0.010*	21	−0.015*	0.008*	0.023*	19
3 months	−0.015*	0.005*	0.020*	40	−0.041*	0.029*	0.070*	58
6 months	−0.023*	0.005†	0.028*	57	−0.050*	0.038*	0.088*	74
9 months	−0.030*	0.013*	0.043*	87	−0.054*	0.046*	0.100*	84
1 year	−0.028*	0.021*	0.049*	100	−0.060*	0.059*	0.119*	100
Post-revision price drift	−0.022*	0.015*	0.038*		−0.046*	0.049*	0.095*	

(continued on next page)

TABLE 3 (continued)

	Low-Innovation Revisions^{a}				High-Innovation Revisions^{a}			
	Bad News	*Good News*	*Difference*	*% of 1-year Return*	*Bad News*	*Good News*	*Difference*	*% of 1-year Return*
Panel C: All-Stars^{d}								
No. of Forecasts	6,145	10,371			31,153	20,233		
Average Forecast Change^{b}	−0.09	0.09			−0.20	0.13		
Size-Adjusted Returns^{c}								
3 days	−0.007*	0.007*	0.014*	40	−0.015*	0.009*	0.024*	23
3 months	−0.015*	0.010*	0.024*	70	−0.038*	0.025*	0.063*	59
6 months	−0.018*	0.010*	0.028*	81	−0.042*	0.028*	0.071*	67
9 months	−0.017*	0.014*	0.031*	89	−0.048*	0.036*	0.085*	80
1 year	−0.024*	0.011†	0.035*	100	−0.060*	0.045*	0.105*	100
Post-revision price drift	−0.017‡	0.003	0.020*		−0.046*	0.034*	0.080*	
Panel D: Earnings-Estimators^{d}								
No. of Forecasts	3,237	5,972			18,884	10,733		
Average Forecast Change^{b}	−0.09	0.09			−0.20	0.12		
Size-Adjusted Returns^{c}								
3 days	−0.006*	0.007*	0.012*	28	−0.015*	0.009*	0.024*	19
3 months	−0.018*	0.004	0.022*	50	−0.048*	0.025*	0.073*	57
6 months	−0.029*	0.007	0.036*	84	−0.058*	0.032*	0.089*	70
9 months	−0.032*	0.010	0.042*	97	−0.064*	0.037*	0.100*	78
1 year	−0.038*	0.006	0.043*	100	−0.081*	0.047*	0.128*	100
Post-revision price drift	−0.032*	−0.002	0.030*		−0.066*	0.036*	0.102*	

(continued on next page)

TABLE 3 (continued)

	Low-Innovation Revisions[a]				High-Innovation Revisions[a]			
	Bad News	*Good News*	*Difference*	*% of 1-year Return*	*Bad News*	*Good News*	*Difference*	*% of 1-year Return*
Panel E: Low Analyst Coverage[e]								
No. of Forecasts	3,633	6,133			22,728	13,906		
Average Forecast Change[b]	−0.06	0.06			−0.20	0.09		
Size-Adjusted Returns[c]								
3 days	−0.003†	0.006*	0.009*	11	−0.016*	0.010*	0.027*	14
3 months	−0.018*	0.009‡	0.027*	32	−0.055*	0.049*	0.105*	55
6 months	−0.024*	0.013†	0.037*	43	−0.073*	0.070*	0.143*	75
9 months	−0.027*	0.032*	0.058*	69	−0.078*	0.088*	0.166*	88
1 year	−0.028‡	0.057*	0.084*	100	−0.081*	0.108*	0.189*	100
Post-revision price drift	−0.025†	0.048*	0.073*		−0.065*	0.096*	0.162*	
Panel F: High Analyst Coverage[e]								
No. of Forecasts	28,509	47,899			148,233	92,522		
Average Forecast Change[b]	−0.08	0.08			−0.19	0.10		
Size-Adjusted Returns[c]								
3 days	−0.006*	0.005*	0.011*	27	−0.015*	0.008*	0.023*	21
3 months	−0.015*	0.005*	0.020*	48	−0.039*	0.025*	0.064*	59
6 months	−0.022*	0.006*	0.028*	68	−0.046*	0.031*	0.076*	71
9 months	−0.027*	0.011*	0.038*	93	−0.051*	0.037*	0.088*	81
1 year	−0.027*	0.014*	0.041*	100	−0.059*	0.048*	0.108*	100
Post-revision price drift	−0.020*	0.008‡	0.028*		−0.045*	0.039*	0.084*	

TABLE 3 (continued)

Two-tailed t-test of difference from zero is significant at: * $\alpha < 0.001$ ‡ $\alpha < 0.01$ † $\alpha < 0.05$

[a] Low-Innovation Revisions are new forecasts that are between the analyst's own prior forecast and the prior consensus. The classification of these signals as good or bad news is based on the analyst's own prior forecast. High-Innovation Revisions are new forecasts that are either above ("Good News") or below ("Bad News") both the prior consensus and the analyst's own prior forecast.

[b] Forecast Change is the new forecast less the analyst's own prior forecast.

[c] Size-Adjusted Returns are the buy-and-hold return for the firm minus the buy-and-hold return for an equal-weighted portfolio of firms in the same firm-size decile. Return accumulation periods are three days (day −1 to +1), three months (day −1 to +64), six months (day −1 to +127), nine months (day −1 to +190), and one year (day −1 to +253), where days are trading days and day 0 is the day of the forecast revision. The post-revision price drift is measured from day +2 to day +253.

[d] *All-Stars* are analysts that ranked first, second, third, or runner-up in their industry in the most recent annual *Institutional Investor* rankings available prior to the revision. *Earnings-Estimators* are analysts ranked first through fifth in their industry in the most recent annual *Wall Street Journal* analyst rankings. Unranked analysts are not classified as either *All-Stars* or *Earnings-Estimators*.

[e] Low (High) Analyst Coverage indicates that the number of analysts covering the firm is below (above) the median analyst coverage for all firms in a given year.

a smaller proportion of the total one-year price response for low coverage firms (11 to 14 percent) than for high coverage firms (21 to 27 percent). This result is weakly significant for both levels of revision innovation (p-values < 0.075).

To summarize, the evidence thus far supports the view that price adjustment is slower and less complete for high-innovation revisions and low-coverage firms. We also find limited evidence that analyst ranking can affect price discovery. However, tabular classifications do not control for the magnitude of the revision signal, or for other firm characteristics that can affect expected returns. In the next section, we conduct more formal statistical tests using multiple regressions.

Revision-Level Regressions

Our multiple regression analysis controls for two measures of the magnitude of the revision signal. Revise measures the size and direction of the forecast revision. To compute this variable, we subtract an analyst's own prior forecast from the new forecast, and express the difference as a percentage of the closing price on day $t - 1$. As an alternative control for the information content of a forecast revision, we also include Ret, computed as the abnormal size-adjusted return during the three-day window centered on the revision date.

To capture the effect of the level of innovation in the revision signal, we define a categorical variable: Signal $= +1$ for high-innovation good news, 0 for low-innovation revision news, and -1 for high-innovation bad news. After we have controlled for both Revise and Ret in the regression, Signal captures the incremental effect of revision innovation (a qualitative feature of the revision) on future returns.

To assess the effect of the analyst's ranking, we group analysts into three categories: *Non-Ranked*, *II All-Stars*, and *WJS Earning-Estimators*. Non-ranked analysts include all analysts not ranked in either survey. *All-Stars* issued 19 percent of the total revisions. Because the *WSJ* groups firms into fewer industries, *Earnings-Estimators* issued only approximately 10 percent of the revisions. The overlap in award winners between the two publications is significant, but not overwhelming—23 percent of the *All-Stars* are also *Earnings-Estimators*. In our multiple regressions, we use two indicator variables (All-Star and Earnings-Estimator) that assume a value of 1 if an analyst is ranked by the respective publication, and 0 otherwise.

We also include an indicator variable to capture the level of analyst coverage a firm receives (Coverage). This variable equals 1 if a firm is followed by more than the median number of analysts, and 0 otherwise. We determine the number of analysts annually, based on the maximum making a one-year-ahead forecast for a given firm in a given year.

Finally, we control for the firm's market-adjusted returns over the prior six months (Momentum), the natural log of a firm's market capitalization at the end of the last calendar year (Size), and its book-to-market ratio as of the end of the last calendar year (B/M). We obtained the financial data for this analysis from the Compustat Price, Dividend, and Earnings file, and the returns data from CRSP. Use of these data reduced the sample size to 280,221 forecast revisions in the most restrictive multiple regression tests. To limit the effect of outliers, we winsorized the extreme top and bottom 1 percent of the B/M, Size, and Revise variables.

Table 4 reports the pairwise Pearson correlation between these variables and two measures of future size-adjusted returns: over the next 6 months (Ret6) and over the next 12 months (Ret12). The statistical significance of these correlations is inflated due to sample dependence problems (addressed in later tests using various statistical corrections). Nevertheless, the magnitude of the correlations generally confirms prior findings. Specifically,

TABLE 4
Pearson Pairwise Correlations among Firm Characteristics, Analyst Forecast Revisions, and Future Returns

	Revise	*Ret*	*Signal*	*Coverage*	*All-Stars*	*Earnings-Estimators*	*Momentum*	*Size*	*B/M*	*Ret6*	*Ret12*
Revise	100.0%	13.8%*	41.7%*	11.5%*	3.0%*	0.9%*	31.0%*	17.1%*	−9.3%*	6.4%*	4.6%*
Ret		100.0%	16.3%*	−0.2%	0.3%	0.0%	13.1%*	1.1%*	1.6%*	1.6%*	3.1%*
Signal			100.0%	0.4%†	1.1%*	−1.1%*	42.0%*	4.2%*	−9.3%*	9.6%*	8.2%*
Coverage				100.0%	7.2%*	3.8%*	−2.3%*	49.1%*	−8.5%*	0.4%†	−0.2%
All-Stars					100.0%	15.7%*	0.2%	12.1%*	2.1%*	0.2%	−0.4%†
Earnings-Est.						100.0%	−1.0%*	6.8%*	0.1%	−0.4%	−0.6%‡
Momentum							100.0%	−2.0%*	−4.6%*	12.0%*	10.5%*
Size								100.0%	−21.8%*	0.2%	−0.8%*
B/M									100.0%	−0.6%*	−1.1%*
Ret6										100.0%	67.0%*
Ret12											100.0%

Two-tailed t-test of difference from zero is significant at: * $\alpha < 0.001$ † $\alpha < 0.01$ † $\alpha < 0.05$. Sample size = 280,221.

Revise = the forecast revision relative to the analyst's own prior forecast, as a percentage of price, with the extreme 1 percent of observations winsorized;
Ret = the firm's size-adjusted return over the three-day event window;
Signal = +1 for high-innovation good news, = 0 for low-innovation news, and = −1 for high-innovation bad news;
Coverage = 1 if the firm is covered by more than the median number of analysts computed annually for the sample, and 0 otherwise;
All-Stars = 1 if the analyst is ranked first, second, third, or runner-up in his industry in the *Institutional Investor* annual rankings, and 0 otherwise;
Earnings-Estimators = 1 if the analyst is ranked first through fifth in his industry in the *Wall Street Journal* annual analyst rankings, and 0 otherwise.
Momentum = market-adjusted returns over the prior six months measured as of day t − 2;
Size = the natural log of a firm's market capitalization at end of the last calendar year with the extreme 1 percent of observations winsorized;
B/M = the firm's book-to-market ratio measured as of the end of the last calendar year with the extreme 1 percent of observations winsorized;
Ret6 = the firm's size-adjusted return from day +2 to day +127, where days are trading days and day 0 is the day of the forecast revision; and
Ret12 = the firm's size-adjusted return from day +2 to day +253, where days are trading days and day 0 is the day of the forecast revision.

in our sample future returns are positively correlated with past price momentum (Momentum). Neither the book-to-market ratio (B/M) nor the firm size (Size) variable is strongly correlated with future returns during this period.

As expected, Signal and Revise are positively correlated (41.7 percent). Both are also highly correlated with Momentum (42.0 percent for Signal and 31.0 percent for Revise), and with Ret, the event-window return (16.3 percent for Signal and 13.8 percent for Revise). Coverage is also highly positively correlated with Size (49.1 percent). Finally, the *II* All-Star and *WSJ* Earnings-Estimator ranking variables are positively correlated, but only at 15.7 percent. In short, these pairwise correlations appear reasonable, and none seem large enough to create multicollinearity problems in the regressions.[17]

The post-revision price drift effect is clearly evident in Table 4. Both measures of future returns are positively correlated with the magnitude of the revision (Revise), the short-window return around the revision (Ret), and the direction of the high-innovation revisions (Signal). Of the three revision variables, Signal appears to have the strongest ability to predict future returns. Over the next 6 to 12 months, Signal has almost the same correlation with future return as the much better known Momentum.

Table 5 reports the result of a pooled cross-sectional regression of future returns on these explanatory variables. The dependent variable is future size-adjusted returns over the next 6 (Ret6) or 12 (Ret12) months. The independent variables include all the conceptual factors we expect to affect the delayed price response, as illustrated in Figure 1: (1) revision quantity—we include both Revise and Ret; (2) revision quality—we include a measure of revision innovation (Signal), as well as two measures of analyst ranking (All-Star and Earnings-Estimator); (3) firm information environment—we include a measure of analyst coverage (Coverage); and (4) other firm characteristics—we include three variables that prior research has shown are related to future returns (Momentum, Size, and B/M). The unit of observation for this analysis is a forecast revision, and the total number of observations is 280,221. As explained earlier, we address the sample dependency problem by computing Huber-White t-statistics.

It is possible that the market's ability to fully appreciate the differential impact of high-innovation revisions vs. low-innovation revisions depends on analyst ranking or analyst coverage. To evaluate the interactions between these effects, we include Coverage × Signal, All-Star × Signal, and Earnings-Estimator × Signal. Model 1 incorporates only revision attributes, Revise, Ret, and Signal. Model 2 adds controls for price momentum, firm size, and the book-to-market ratio. Finally, Model 3 incorporates all the explanatory variables, including the interaction terms.

The Signal variable construction implicitly assumes a long position in a firm for each high-innovation good news revision, and a short position for each bad news revision. Therefore, the estimated coefficient for Signal represents the average abnormal return from a single bet (either a long or a short position) made on each high-innovation revision. To compute the average *hedge* return—i.e., the combination of a long position with a short position—we double this coefficient.[18]

This approach is analogous to the post-earnings-announcement literature, which focuses on quarterly Standardized Unexpected Earnings (SUE) as the unit of observation. When a

---

[17] We also conduct more formal tests of collinearity on an unreported regression with all of the independent variables and no interaction terms. The highest condition index for any variable was 2.00 (index values above 30 would indicate a problem) and the highest variance inflation factor was 1.42 (factors below 10 are generally not considered a problem)

[18] Following prior studies, we use the term *hedge* to imply cash-flow neutrality rather than zero risk exposure.

TABLE 5
Panel Regression of Future Size-Adjusted Returns on *Ex Ante* Firm and Forecast Revision Characteristics

		Ret6			*Ret12*	
	Model 1	*Model 2*	*Model 3*	*Model 1*	*Model 2*	*Model 3*
Intercept	−0.008†	−0.013	−0.011	−0.010	0.014	0.016
	(−2.41)	*(−1.05)*	*(−0.90)*	*(−1.19)*	*(0.53)*	*(0.62)*
Revise	0.661*	0.325	0.236	0.512	0.033	−0.075
	(3.59)	*(1.67)*	*(1.20)*	*(1.55)*	*(0.09)*	*(−0.20)*
Ret	−0.008	−0.034	−0.033	0.152‡	0.112†	0.114†
	(−0.36)	*(−1.54)*	*(−1.50)*	*(3.20)*	*(2.45)*	*(2.48)*
Signal	0.032*	0.020*	0.045*	0.049*	0.030*	0.060*
	(9.57)	*(7.39)*	*(8.01)*	*(6.77)*	*(5.55)*	*(5.42)*
Coverage			0.001			0.003
			(0.20)			*(0.17)*
Coverage × Signal			−0.027*			−0.034‡
			(−4.55)			*(−2.76)*
All-Star			0.000			−0.002
			(−0.18)			*(−0.95)*
All-Star × Signal			−0.008‡			−0.007
			(−3.17)			*(−1.46)*
Earnings-Estimator			0.000			−0.001
			(−0.45)			*(−0.57)*
Earnings-Estimator × Signal			0.008‡			0.014†
			(2.60)			*(2.39)*
Momentum		0.093*	0.093*		0.147*	0.146*
		(6.07)	*(6.03)*		*(4.29)*	*(4.27)*
Size		0.000	0.000		−0.003	−0.004
		(0.02)	*(−0.20)*		*(−0.85)*	*(−0.82)*
B/M		0.006	0.005		−0.010	−0.010
		(0.47)	*(0.44)*		*(−0.39)*	*(−0.40)*
Adjusted R^2	0.010	0.017	0.018	0.007	0.013	0.013

Two-tailed t-test of difference from zero is significant at: * $\alpha < 0.001$ ‡ $\alpha < 0.01$ † $\alpha < 0.05$.

We use the Huber-White maximum likelihood estimation procedure to correct for sample dependencies among multiple revisions for the same firm. Huber-White t-statistics are italicized and in parentheses.

Sample size = 280,221.

Variable definitions appear in the footnotes to Table 4.

firm has a positive SUE quarter, followed by a negative SUE quarter, the reported return for the hedge strategy would reflect a long position in quarter-one and a short position in quarter-two. To the extent that the two positions have overlapping holding periods for future returns, they cancel each other out, resulting in zero hedge returns for the duration of the overlap.

Results in Table 5 confirm the relative importance of revision innovation in explaining the post-revision price drift. As expected, the model has modest overall explanatory power for future returns. However, the Huber-White t-statistic on Signal is 8.01 for the Ret6 regression (Model 3), and Signal has the highest t-statistic in all six regressions. Table 5 shows that a hedge strategy based on Signal yields 10 percent abnormal returns over the next year after controlling for the magnitude of the revision signal and the three-day abnormal return over the event window (Model 1). The same strategy yields 6 percent over the next year after controlling for Revise, Ret, Momentum, Size, and B/M (Model 2).

The Revise and Ret variables compete for explanatory power. When we exclude one, the other is positively associated with future returns over both 6- and 12-month horizons. However, when we include both variables, Revise is marginally better at predicting 6-month returns, while Ret is marginally better at predicting 12-month returns. More importantly, neither Ret nor Revise is as important as Signal in terms of their ability to explain future returns. These results suggest that Signal captures *qualitative* aspects of an analyst's forecast revision that helps explain cross-sectional variations in the post-revision price drift.

The main effects of analyst coverage (Coverage) and the two analyst ranking variables (All-Star and Earnings-Estimator) are not statistically significant (Model 3). These findings show that, after controlling for the other variables, analyst coverage and the analyst's ranking, *per se*, do not explain significant variation in future returns.

More interesting are the interaction results. The estimated coefficient for Coverage × Signal is reliably negative, indicating that future returns to Signal are lower for firms with higher analyst coverage. Similarly, the estimated coefficient for All-Star × Signal is also reliably negative for the Ret6 regressions. This result suggests that returns over the next six months to Signal are reliably lower when the analyst is an All-Star. In contrast, the estimated coefficient for Earnings-Estimator × Signal is reliably positive for both Ret6 and Ret12, suggesting that future returns to Signal are higher when the analyst is an Earnings-Estimator.[19]

These results suggest that the market impounds the information in innovative forecast revisions more quickly when these revisions are issued by an All-Star analyst or are issued for more heavily covered firms. Conversely, the market is slower to impound the information in innovative forecast revisions when these revisions are issued by an Earnings-Estimator or are issued for firms with lower analyst coverage.

In Model 3, the inclusion of the interaction terms increases the estimated coefficient for Signal. Because the interaction terms remove the incremental effects of ranked analysts and heavily covered firms, the coefficient on Signal in Model 3 reflects the post-revision price drift associated with revisions by unranked analysts in lightly covered firms. Our results show that the post-revision drift is higher for these forecast revisions.

To summarize, Table 5 regressions show that the level of innovation in the revision signal (Signal) is an important factor in explaining cross-sectional variations in the post-revision price drift. Our results suggest that the other primary variables of interest (Coverage, All-Star, and Earnings-Estimator) affect future returns by attenuating or exacerbating the market's ability to interpret the information in Signal. Specifically, the post-revision price drift is more pronounced for innovative revisions issued by *Earnings-Estimators* in low-coverage firms, and is less pronounced for revisions issued by *All-Stars* for high-coverage firms.

---

[19] We also included additional two-way interaction terms between analyst coverage and analyst ranking (Coverage × Earnings-Estimator and Coverage × All-Star) and three-way interaction terms (Coverage × Signal × Earnings-Estimator and Coverage × Signal × All-Star). These variables are not incrementally important in the regression, and do not change our results significantly.

Subsequent Earnings-Related News Announcements

To better understand the nature of the post-revision price drift, we conduct further tests around subsequent earnings-related news announcement dates. Specifically, we examine abnormal returns in three-day windows centered on subsequent earnings announcements and forecast revisions. The main purpose of these tests is to discriminate between the market-inefficiency and risk-based explanations for the post-revision price drift. If the post-revision drift is due to a delayed market response to current information about future earnings *and* this misperception is corrected when further information about future earnings is released, then subsequent abnormal returns should cluster around the future announcements of earnings news. Conversely, if the price drift is due to omitted risk variables, we should not observe higher abnormal returns concentrated around the release of subsequent earnings news.

Table 6 reports the results of these tests. Earlier results show that the post-revision price drift is associated primarily with high-innovation forecasts. Therefore, this analysis focuses only on the high-innovation revisions. Table values represent mean size-adjusted returns over three-day windows centered on the next four quarterly earnings announcements and the next six analyst forecast revisions. We include only revisions that occurred within 12 months of the original revision. To eliminate revisions that cluster closely together in time, we exclude from this analysis any revision that occurs within the three-day announcement window of a revision that is already included in the sample.

Table 6 reports returns for good-news and bad-news revisions, as well as the difference between the two (i.e., the cash-neutral hedge return). We also report the percentage of the abnormal one-year return to each hedge strategy that is realized in each three-day event window. Panel A presents the results for all firms, Panel B presents the results for high analyst coverage firms, and Panel C presents the results for low analyst coverage firms.

Panel A of Table 6 shows that the hedge return is positive around all four subsequent earnings announcements. Collectively, 11.1 percent of the abnormal return realized over the next 12 months is earned in the days immediately surrounding the next four earnings releases. Assuming expected returns do not vary daily, we expect 4.7 percent (12/253) of the abnormal return to occur over 12 trading days. The higher proportion of abnormal returns observed around earnings announcements suggests that at least a portion of the delayed price response is due to misperception about future earnings, which is corrected around future earnings release dates.

Even more striking are the results related to the next six analyst forecast revisions. These results show that the market continues to be surprised in a negative direction for bad-news revisions and in a positive direction for good-news revisions, when other analysts subsequently revise their forecasts. In fact, we observe a significant price reaction over at least the next *six* forecast revisions. Collectively, the abnormal returns around the next six revisions represent 26.8 percent of the total one-year abnormal returns to the hedge strategy—compared to 7.1 percent (18/253) by chance alone. This evidence suggests that much of the incomplete price response to a given revision is corrected when later analysts revise their forecasts.

The next two panels of Table 6 show that this result holds for both high-coverage and low-coverage firms. For low-coverage firms (Panel C), 13.9 percent of the price drift occurs around the next four earnings announcements, and 15.1 percent of the price drift occurs around the next six forecast revisions. For high-coverage firms (Panel B), where the drift itself is lower, a lower percentage of that drift (10.5 percent) occurs around future earnings announcements while a higher proportion of the drift (30.2 percent) appears around the next six forecast revisions.

TABLE 6

Mean Size-Adjusted Returns for Three-Day Windows around Subsequent Earnings Announcements and Subsequent Analyst Forecast Revisions for High-Innovation Signals[a,b]

	Subsequent Earnings Announcements				*Subsequent Analyst Forecast Revisions*			
	Bad News	*Good News*	*Difference*	*% of 1-Year Return*	*Bad News*	*Good News*	*Difference*	*% of 1-Year Return*
Panel A: All Firms								
1st	0.000	0.006*	0.006*	6.3	−0.007*	0.003*	0.010*	10.2
2nd	0.004*	0.006*	0.002*	2.1	−0.003*	0.001*	0.004*	4.3
3rd	0.005*	0.006*	0.001*	1.4	−0.003*	0.000	0.003*	3.1
4th	0.004*	0.005*	0.001*	1.3	−0.004*	0.000	0.003*	3.5
5th					−0.003*	0.000	0.003*	3.3
6th					−0.003*	−0.001‡	0.002*	2.4
Panel B: High Analyst Coverage[c]								
1st	0.001*	0.006*	0.005*	5.8	−0.007*	0.003*	0.010*	11.3
2nd	0.005*	0.006*	0.002*	1.8	−0.003*	0.001*	0.004*	4.8
3rd	0.005*	0.006*	0.001*	1.4	−0.003*	0.000	0.003*	3.6
4th	0.004*	0.005*	0.001*	1.5	−0.003*	0.000	0.003*	3.9
5th					−0.003*	0.000	0.003*	3.7
6th					−0.003*	0.000†	0.003*	2.9

(*continued on next page*)

TABLE 6 (continued)

	Subsequent Earnings Announcements				*Subsequent Analyst Forecast Revisions*			
	Bad News	*Good News*	*Difference*	*% of 1-Year Return*	*Bad News*	*Good News*	*Difference*	*% of 1-Year Return*
Panel C: Low Analyst Coverage[c]								
1st	−0.004*	0.010*	0.014*	8.4	−0.008*	0.002‡	0.010*	6.0
2nd	0.002‡	0.008*	0.006*	3.4	−0.004*	0.001	0.005*	2.9
3rd	0.004*	0.006*	0.002†	1.5	−0.004*	−0.002†	0.002†	1.5
4th	0.003*	0.004*	0.001	0.6	−0.005*	−0.001	0.004*	2.4
5th					−0.005*	−0.002	0.003*	2.1
6th					−0.003*	−0.003*	0.000	0.2

Two-tailed t-test of difference from zero is significant at: * $\alpha < 0\ .001$ ‡ $\alpha < 0.01$ † $\alpha < 0.05$.

[a] For each revision, we report the mean size-adjusted returns earned around the next four earnings announcements and the next six forecast revisions, as well as the percentage of the abnormal one-year hedge return realized in each event window. We include only revisions that occur within 12 months of the original revision, and exclude any revision that takes place within three days of a prior revision. Size-adjusted returns are buy-and-hold returns for three-day windows centered on the earnings announcement or forecast revision date, minus the return for an equal-weighted portfolio of firms in the same firm-size decile.

[b] High-Innovation Signals are forecast revisions that are either above ("Good News") or below ("Bad News") both the prior consensus and the analyst's own prior forecast.

[c] Low Analyst Coverage indicates that the number of analysts covering the firm is below the median analyst coverage among all firms in a given year. High Analyst Coverage indicates the number of analysts is above the median.

Firm-Level Analysis

Because our primary focus is on understanding how the market assimilates the information content of individual revisions, rather than on trading profits, our analyses thus far have used individual analyst forecast revisions as the unit of observation. However, in Table 7, we also report the results of a cross-sectional regression of future returns on firm-level explanatory variables. To perform this analysis, we convert revision-level information into firm-level indicator variables, and regress future size-adjusted returns on these variables.

We have two purposes for conducting this test. First, firm-level analysis provides an alternative solution to the sample dependence problem arising from including multiple observations for the same firm. The regressions reported in Table 5 address this econometric challenge using Huber-White estimation techniques. The firm-level regressions reported in Table 7 entirely eliminate revision-level dependency problems by ensuring that each firm appears only once in each regression. Because these results are based on monthly cross-sectional (i.e., Fama-MacBeth type) regressions, the Huber-White estimation technique is unnecessary. However, due to overlapping holding periods, the time-series t-statistics for the monthly cross-sectional coefficient estimates can still be inflated due to positive serial correlation. We adjust for this problem using Newey-West procedure (Newey and West 1987) procedures.

Second, Table 7 provides information on the economic significance of revision-based investment strategies, assuming monthly rebalancing. The results reported in this table are not exactly the same as those one would obtain from a revision-level strategy, because this test does not capture the returns that accrue between day $t + 1$ and the end of the first calendar month. Nevertheless, this table provides some insight on how an investor might fare if he or she were to trade on a revision-level signal using firm-level information at the end of each calendar month.

To construct this table, we perform monthly cross-sectional firm-level regressions. The dependent variable is the size-adjusted return for the six months (Ret6) or the 12 months (Ret12) following the month of the forecast revision. The key independent variable, FSignal, is a firm-level indicator variable that captures the information content of the most recent forecast revision for that firm. Specifically, FSignal $= +1$ if the most recent forecast revision issued in the month prior to portfolio formation is high-innovation and upward, -1 if the revision is high innovation and downward, and 0 otherwise.[20]

We also include three interaction terms: (1) the Coverage × FSignal variable equals FSignal multiplied by an indicator variable that equals 1 if the firm is in the high-coverage group, 0 otherwise; (2) the All-Star × FSignal variable equals FSignal multiplied by an indicator variable that equals 1 if the analyst is an All-Star, 0 otherwise; and, (3) the Earnings-Estimator × FSignal variable is equal to FSignal multiplied by an indicator variable that equals 1 if the analyst is an Earnings-Estimator, 0 otherwise. Analogous to Table 5, these firm-level interaction terms allow us to assess the incremental effect of analyst coverage and the analyst's ranking on post-revision price drift. Finally, we also control for Revise and Ret based on the most recent high-innovation revision, and three firm characteristics (Momentum, Size, and B/M).

Although our conversion from revision-level information to firm-level information involves significant loss of information, the results reported in Table 7 generally confirm the revision-level findings in Table 5. Specifically, even after controlling for Size, B/M, and Momentum, the sign of the most recent high-innovation earnings forecast revision in the prior month continues to predict future returns. The coefficients on FSignal provide an estimate of the returns from cash-neutral hedge strategies involving unranked analysts in low-coverage firms: annualized returns range from 15.2 percent based on the Ret6 regression (3.8 percent × 4) to 11.0 percent based on the Ret12 regression (5.5 percent × 2).

---

[20] One could convert the revision-level signal to a firm-level signal in many ways, but each involves some loss of information. In earlier versions of this paper, we reported results based on alternative definitions of the Signal variable including: (1) the sign of the change in the monthly consensus forecast; (2) the net number of upward vs. downward revisions over the past month; and, (3) the net number of high-innovation upward vs. downward revisions over the past month. The results are all quite similar. We report results based on only the most recent forecast revision, because it seems most faithful to the spirit of the revision-level analysis we conducted earlier.

The coefficients on Coverage × FSignal show that, controlling for firm size, returns to an FSignal strategy are reliably lower for firms with high analyst coverage. Similarly, the

TABLE 7

Cross-Sectional Regressions of Future Size-Adjusted Returns on Firm-Level Explanatory Variables[a]

	Ret6	*Ret12*
Intercept	−0.011†	−0.016†
	(−2.22)	*(−2.03)*
Revise	0.011*	0.020*
	(4.28)	*(6.02)*
Ret	0.004†	0.009‡
	(2.55)	*(2.67)*
FSignal	0.038*	0.055*
	(7.74)	*(6.82)*
Coverage × FSignal	−0.025*	−0.034*
	(−4.54)	*(−5.99)*
All-Star × FSignal	−0.006†	−0.012‡
	(−2.45)	*(−2.81)*
Earnings-Estimator × FSignal	0.002	0.005
	(0.32)	*(0.49)*
Momentum	0.027*	0.032*
	(5.06)	*(3.70)*
Size	−0.008‡	−0.026*
	(−3.29)	*(−4.96)*
B/M	0.008	0.009
	(1.31)	*(0.78)*
R^2	0.043	0.033

Two-tailed t-test of difference from zero is significant at: * $\alpha < 0.001$ ‡ $\alpha < 0.01$ † $\alpha < 0.05$.

[a] This table reports the results of cross-sectional regressions of future size-adjusted returns on firm-level explanatory variables. The reported coefficient estimates and adjusted R^2s are the time-series averages from 63 monthly cross-sectional regressions. All reported t-statistics are corrected for autocorrelation using the Newey-West (1987) algorithm. Firms are included in a monthly regression only if they have at least one analyst forecast revision during the month. The number of firms per month in each regression ranges from 880 to 2,182.

The dependent variable for each regression is either the size-adjusted return for the 6 months following the month of the forecast revision (Ret6), or the size-adjusted return for the 12 months following the month of the forecast revision (Ret12).

(*continued on next page*)

TABLE 7 (continued)

The independent variables are:

Revise = the change in the monthly consensus forecast in the month prior to portfolio formation, expressed as a percentage of price;
Ret = the firm's size-adjusted return over the three-day event window around the last forecast issued in the month prior to portfolio formation;
FSignal = 1 for high-innovation good news, −1 for high-innovation bad news, and 0 otherwise; based on the last forecast issued in the month prior to portfolio formulation;
Coverage = 1 if the firm is covered by more than the median number of analysts where the median number of analysts is computed annually for the sample, and 0 otherwise;
All-Star = 1 if the analyst is ranked first, second, third, or runner-up in his industry in the *Institutional Investor* annual rankings, and 0 otherwise;
Earnings Estimator = 1 if the analyst is ranked first through fifth in his industry in the *Wall Street Journal* annual analyst rankings, and 0 otherwise;
Momentum = market-adjusted returns over the prior six months measured as of day t − 2. The variable is scaled by subtracting the mean and dividing by the standard deviation;
Size = the natural log of a firm's market capitalization measured as of the end of the last calendar year, with the extreme 1 percent of the observations winsorized. The variable is scaled by subtracting the mean and dividing by the standard deviation; and
B/M = the firm's book-to-market ratio measured as of the end of the last calendar year, with the extreme 1 percent of the observations winsorized. The variable is scaled by subtracting the mean and dividing by the standard deviation.

coefficients on All-Star × FSignal show that returns to an FSignal strategy are reliably lower for revisions issued by *All-Stars*. The coefficient on Earnings-Estimator × FSignal is insignificant, perhaps because relatively few high-innovation revisions in this firm-level analysis are issued by *WSJ Earnings-Estimators*. Similar to Table 5, FSignal has higher t-statistics than either Ret or Revise. These findings confirm that our prior inferences are not driven by the use of multiple revisions from the same firm.

V. CONCLUSION

Much capital market research in accounting over the past 20 years has assumed that the price adjustment to information is instantaneous. This basic assumption has influenced the way we select research topics, design empirical tests, and interpret research findings. In this study, we provide evidence that price discovery, at least as it pertains to analyst forecast revisions, is a protracted and complex process that deserves more scholarly attention.

Using a large sample of individual forecast revisions, we conjecture, and show, that the magnitude of the post-revision price drift is a function of various revision and firm characteristics. First, analogous to findings in the post-earnings-announcement drift literature, we show that the delayed price response is a function of *revision quantity*—that is, firms with good-news revisions of a higher magnitude experience more positive subsequent returns and firms with bad-news revisions experience negative subsequent returns.

Second, we find that several measures of *revision quality* also help explain the delayed price response. Specifically, we show that the market does not make a sufficient distinction between analysts who bring new information to market (high-innovation revisions) and those who may be simply herding toward the consensus forecast (low-innovation revisions). We also find that an analyst's ranking affects the speed and efficacy with which prices reflect the information content of a revision. Specifically, we show that, compared to unranked analysts, both *II All-Stars* and *WSJ Earnings-Estimators* elicit a stronger immediate market reaction. However, the post-revision drift is significantly smaller for *All-Stars* and significantly larger for *Earnings-Estimators*. We attribute this difference to the celebrity status of the *All-Stars*, and the fact that *Earnings-Estimators* are equally accurate but often more obscure analysts.

Third, we show that the delayed price response to forecast revisions is reliably smaller for firms covered by more analysts. Controlling for other variables, the number of analysts itself does not predict future returns. However, higher analyst coverage leads to faster and more complete assimilation of the information conveyed by the level of revision innovation.

Finally, we find that a substantial portion of the delayed price reaction occurs around subsequent earnings-news event dates. This result is not due to revisions that are closely clustered in time. After excluding all revisions that occur within a three-day window of any prior revision, we find that a disproportionate amount of the post-revision price drift occurs within short-windows around the next *four* earnings announcements and the next *six* forecast revisions. This evidence suggests that the market is adjusting slowly to the information contained in individual forecast revisions, and that future revisions by other analysts act as catalysts in the price discovery process.

In sum, the market seems to overlook certain qualitative aspects of earnings revisions. After controlling for the sign and magnitude of the forecast revision itself, as well as various firm characteristics known to be associated with expected returns, the market still seems to underappreciate more subtle aspects of the revision, such as the level of innovation in the revision signal, and whether the analyst is a superior forecaster from a smaller brokerage firm (*WSJ Earnings-Estimators*). On the other hand, the market seems to assimilate earnings information faster and more completely when more analysts are following the firm and when the analyst issuing the forecast has celebrity status (*II All-Stars*). These findings help us to better understand the price discovery mechanism and the boundaries of market efficiency.

Another interpretation of our results is that good-news firms (i.e., firms whose earnings have just been revised upward) are riskier than bad-news firms, and that the higher average returns earned by good-news firms represent compensation for this higher level of risk. This interpretation allows us to remain within the comfortable confines of the efficient market hypothesis. However, it appears inconsistent with several aspects of our evidence. First, it leaves open questions as to why firm-specific risk should be *positively* correlated with the release of good news about earnings.[21] Second, it begs the question of what risk factor might be missing even after we have controlled for price momentum, firm size, and book-to-market. Third, it does not explain why a large portion of the abnormal returns occurs around the dates of subsequent earnings news releases. Finally, it does not explain why the magnitude of the abnormal returns should be a function of signal quality, the analyst's ranking, and the extent of analyst coverage. Given our collective evidence, we are skeptical of the risk-based explanation.

Finally, our results are related to the current debate over Regulation Fair Disclosure (FD). This regulation, effective October 2001, prohibits management from discussing company prospects with an analyst unless the same information is simultaneously conveyed to others, including other analysts. Our evidence suggests that individual financial analysts bring relevant information to market, and that multiple analysts enhance the speed of the price discovery process. One cannot help but wonder whether this will continue to hold

---

[21] According to the risk explanation, firms that earn higher subsequent returns are riskier in some unknown dimension. Yet in this study, firms with positive rather than negative news earn higher future returns. Therefore, a risk-based interpretation would need to explain why good-news firms are becoming riskier.

APPENDIX

Sample Screen Capture of First Call® Information Delivered by FactSet®

Corning Inc. (GLW) Individual Analyst Earnings Forecast Revisions for Fiscal Year Ended 12/2001

CORNING INC		11-Jan-2002
GLW	21935010	NYSE
United States: MULTI-IND CAP GOOD		USD
Division Factor: 1.02		*Hold (2.77)*

**Denotes excluded by I/B/E/S*

Date of Latest Change	*Broker*	*Analyst*	*12/2001*	*Change*		*12/2002*	*Change*	
				Amount	*%*		*Amount*	*%*
11-Jan-2002	**Mean (23 Ests)**		**.44**			**NA**		
28-Nov-2001	Needham & Company	LANGLEY,M	.44			−.37		
20-Nov-2001	Thomas Weisel Partners	BUNTING,J	.45	−.46	−50.5	−.39	−1.56	−133.3
14-Nov-2001	Roth Capital Partners	KANG,D	.45			−.45		
06-Nov-2001	Morgan Stanley, Dean Witter Dis	JACKSON,D	.45	−.15	−25.0	−.31	−.35	−875.0
24-Oct-2001	J.P. Morgan	LIPTON,J	.41	−.14	−25.5	−.05	−.19	−135.7
24-Oct-2001	Davenport & Co. Of Virginia, In	JOHNSTONE,D	.44	−.07	−13.7	−.41	−.41	
24-Oct-2001	Merrill Lynch	FOX,S	.43	−.07	−14.0	−.17	−.02	
24-Oct-2001	C.E. Unterberg, Towbin	DRAGONE,J	.42	−.10	−19.2	−.40	−.28	
23-Oct-2001	Goldman Sachs & Co.	SUNDARAM,N	.45	.01	2.3	−.31	−.35	−875.0
23-Oct-2001	Dresdner Kleinwort Benson	KINGSTON,S	.48			−.37	−.26	
23-Oct-2001	First Union Securities	KOFFLER,S	.42	−.17	−28.8	−.47	−.54	−771.4
23-Oct-2001	Frost Securities	HAIDER,S	.45	−.02	−4.3	−.45	−.95	−190.0
23-Oct-2001	Deutsche Banc Alex. Brown-Us	SRIKANTH,R	.43	−.12	−21.8	−.55	−.57	−2850.0
22-Oct-2001	Robertson Stephens	VEERAPPANA,A	.42	−.18	−30.0	−.48	−.61	−469.2
22-Oct-2001	A. G. Edwards & Sons, Inc.	ANDREW,P	.44	−.15	−25.4	−.36	−.57	−271.4
22-Oct-2001	Banc of America Securities Llc.	CRESPI,C	.46	−.12	−20.7	−.44	−.63	−331.6
22-Oct-2001	Credit Lyonnais Securities (USA	LOWY,G	.42	−.47	−52.8	−.36	−.09	
22-Oct-2001	UBS Warburg	THEODOSOPOULC	.45	−.14	−23.7	−.40	−.55	−366.7
22-Oct-2001	Wit Soundview	SLOCUM,K	.45	−.02	−4.3	−.10	−.15	−300.0
22-Oct-2001	William Blair & Company, L.L.C.	TANGO,R	.47	−.07	−13.0	−.44	−.54	−540.0
19-Oct-2001	Salomon Smith Barney Inc	ANDERSON,T	.44	−.02	−4.3	−.35	−.20	
19-Oct-2001	Credit Suisse First Boston Corp	SCHUETZ,M	.47	−.05	−9.6	−.46	−.20	
19-Oct-2001	ABN Amro Incorporated	POTTER\HOLMSTE	.45	−.05	−10.0	−.25	−.05	

under the new regulations. At first blush, it seems unlikely that the forecast revisions of individual analysts would become more useful under Regulation FD. This appears to be an interesting avenue for future research.

REFERENCES

Ball, R., and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (Autumn): 159–178.

Barber, B., R. Lehavy, M. McNichols, and B. Trueman. 2003. Prophets and losses: Reassessing the returns to analysts' stock recommendations. *Financial Analysts Journal* (forthcoming).

Bernard, V. L., and J. K. Thomas. 1989. Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27 (Supplement): 1–36.

———, and ———. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13 (December): 305–341.

Bhattacharya, N. 2001. Investors' trade size and trading responses around earnings announcements: An empirical investigation. *The Accounting Review* 76 (April): 221–244.

Bloomfield, R., R. Libby, and M. W. Nelson. 2000. Underreactions, overreactions, and moderated confidence. *Journal of Financial Markets* 3 (May): 113–137.

Brennan, M. J., N. Jegadeesh, and B. Swaminathan. 1993. Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies* 6 (Winter): 799–824.

Brown, L., and M. Rozeff. 1980. Analysts can forecast accurately. *Journal of Portfolio Management* 6 (Spring): 31–34.

———. 2001. How important is past analyst forecast accuracy? *Financial Analysts Journal* 57 (December): 44–49.

———, and E. Mohammad. 2001. Profiting from predicting individual analyst earnings forecast accuracy. Working paper, Georgia State University.

Butler, K., and L. Lang. 1991. The forecast accuracy of individual analysts: Evidence of systematic optimism and pessimism. *Journal of Accounting Research* 29 (Spring): 150–156.

Chan, L. K., N. Jegadeesh, and J. Lakonishok. 1996. Momentum strategies. *Journal of Finance* 51 (December): 1681–1713.

Clement, M. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27 (June): 285–303.

———, and S. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78 (January): 227–249.

Diggle, P. J., K. Y. Liang, and S. L. Zeger. 1994. *Analysis of Longitudinal Data.* Oxford, U.K.: Clarendon Press.

Elgers, P. T., M. H. Lo, and R. J. Pfeiffer, Jr. 2001. Delayed security price adjustments to financial analysts' forecasts of annual earnings. *The Accounting Review* 76 (October): 613–632.

Elton, E. J., M. J. Gruber, and M. Gultekin. 1981. Expectations and share prices. *Management Science* 27 (September): 975–987.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (June): 427–465.

Foster, G., C. Olsen, and T. Shevlin. 1984. Earnings releases, anomalies, and the behavior of security returns. *The Accounting Review* 59 (October): 574–603.

Freeman, R. N., and S. Tse. 1989. The multiperiod information content of accounting earnings: Confirmations and contradictions of previous earnings reports. *Journal of Accounting Research* 27 (Supplement): 49–79.

Givoly, D., and J. Lakonishok. 1979. The information content of financial analysts' forecasts of earnings. *Journal of Accounting and Economics* 1 (December): 165–185.

———, and ———. 1980. Financial analysts' forecast of earnings: The value to investors. *Journal of Banking and Finance* 4 (September): 221–233.

Griffin, D., and A. Tversky. 1992. The weighing of evidence and the determinants of confidence. *Cognitive Psychology* 24 (July): 411–435.

Griffin, P. 1976. Competitive information in the stock market: An empirical study of earnings, dividends, and analysts' forecasts. *Journal of Finance* 31 (May): 631–650.

Hong, H., T. Lim, and J. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55 (February): 265–295.

Huber, P. 1967. The behavior of maximum likelihood estimates under nonstandard conditions. In *Procedures of the Fifth Annual Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, edited by L. M. LeCam, and J. Neyman, 221–233. Berkeley, CA: University of California Press.

Imhoff, E., and G. Lobo. 1984. Information content of analysts' composite forecast revisions. *Journal of Accounting Research* 22 (Autumn): 541–554.

Jacob, J., T. Lys, and M. Neale. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28 (November): 51–82.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48 (March): 65–91.

———, J. Kim, S. Krische, and C. Lee. 2001. Analyzing the analysts: When do recommendations add value? Working paper, University of Illinois and Cornell University.

Kothari, S. P. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31 (September): 105–231.

Lee, C. M. C. 2001. Market efficiency and accounting research: A discussion of "Capital market research in accounting" by S. P. Kothari. *Journal of Accounting and Economics* 31 (September): 233–253.

Mikhail, M. B., B. R. Walther, and R. H. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* 35 (Supplement): 131–157.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definitive, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55 (May): 703–708.

O'Brien, P. C. 1990. Forecast accuracy of individual analysts in nine industries. *Journal of Accounting Research* 28 (Autumn): 286–304.

Park, C. W., and E. K. Stice. 2000. Analyst forecasting ability and the stock price reaction to forecast revisions. *Review of Accounting Studies* 5 (September): 259–272.

Sinha, P., L. D. Brown, and S. Das. 1997. A re-examination of financial analysts' differential earnings forecast accuracy. *Contemporary Accounting Research* 14 (Spring): 1–42.

Sloan, R. G. 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (July): 289–315.

Stickel, S. E. 1990. Predicting individual analyst earnings forecasts. *Journal of Accounting Research* 28 (Autumn): 409–417.

———. 1991. Common stock returns surrounding earnings forecast revisions: More puzzling evidence. *The Accounting Review* 66 (April): 402–416.

———. 1992. Reputation and performance among security analysts. *Journal of Finance* 47 (December): 1811–1836.

Walther, B. R. 1997. Investor sophistication and market earnings expectations. *Journal of Accounting Research* 35 (Autumn): 157–179.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (May): 817–838.

Womack, K. 1996. Do brokerage analysts' recommendations have investment value? *Journal of Finance* 51 (March): 137–167.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 227–249

Do Investors Respond to Analysts' Forecast Revisions as if Forecast Accuracy Is All That Matters?

Michael B. Clement
Senyo Y. Tse
The University of Texas at Austin

ABSTRACT: Prior research suggests that investors' response to analyst forecast revisions increases with the analyst's forecast accuracy. We extend this research by examining whether investors appear to extract all of the information that analyst characteristics provide about forecast accuracy. We find that only some of the analyst characteristics that are associated with future forecast accuracy are also associated with return responses to forecast revisions. This suggests that investors fail to extract some of the information that analyst characteristics can provide about future forecast accuracy. In addition, forecast properties other than expected accuracy appear to be value-relevant. For example, investors respond more strongly to forecasts released earlier in the year and to forecasts by analysts employed by large brokerages than appears warranted by the ability of forecast timeliness and broker size to predict forecast accuracy. We conclude that investors respond to analysts' forecast revisions as if forecast accuracy is *not* all that matters.

Keywords: *analyst forecast accuracy; analyst characteristics; security returns.*

Data Availability: *Data are available from public sources.*

I. INTRODUCTION

Theory predicts that investors' response to forecast revisions increases with the expected accuracy of the forecast (Abarbanell et al. 1995), and empirical research documents results

---

We are grateful for helpful comments from Andrew Alford, Mark Bagnoli, Sudipta Basu, Larry Brown, Richard Frankel, Ross Jennings, Jeffrey Miller, Susan Watts, Richard Willis, Hong Xie, two anonymous reviewers, and workshop participants at Baruch College, the University of Chicago, Duke University, Emory University, Georgia State University, Harvard Business School, Michigan State University, the University of Notre Dame, the Ohio State University, Purdue University, Rice University, the University of Texas at Austin, the University of Wisconsin at Madison, the 2000 Analyzing Corporate Earnings Seminar and the 2000 Corporate Earnings Analysis Seminar. Bill Mayew provided valuable research assistance. We also gratefully acknowledge the contribution of I/B/E/S International Inc. for providing earnings per share forecast data, available through the Institutional Brokers Estimate System. These data have been provided as part of a broad academic program to encourage earnings expectation research.

Submitted February 2001
Accepted August 2002

consistent with this prediction (Stickel 1992; Park and Stice 2000; Gleason and Lee 2003). Empirical research also finds an association between forecast accuracy and several forecast and analyst characteristics: forecast timeliness, analyst experience, broker size, and the number of firms and industries the analyst follows (Mikhail et al. 1997; Clement 1999; Jacob et al. 1999).[1] Our study investigates whether *ex ante* predicted forecast accuracy is sufficient to explain differences in investors' response to analyst forecast revisions. If so, then individual analyst characteristics should be associated with returns only to the extent that those characteristics predict forecast accuracy. Implicitly, investors would use all relevant analyst characteristics to predict forecast accuracy, and then would respond to each forecast revision according to the analyst's expected forecast accuracy. Analyst characteristics that are associated with increased forecast accuracy would therefore also be associated with greater return responses to forecast revisions.

There are at least three reasons, however, why the weights investors place on forecasts may diverge from the weights suggested by a model that uses individual analyst characteristics to predict forecast accuracy. First, factors other than forecast accuracy may be value-relevant. For example, investors may respond more strongly to forecasts issued early in the fiscal year (timely forecasts) even though those early forecasts are generally less accurate than later forecasts. As Schipper (1991, 113) notes:

> To the extent having the forecast sooner (even at the cost of less accuracy) implies greater investing profits to consumers of analysts' earnings forecasts, the loss function implied by pleasing customers will create a preference for timeliness.

Second, investors may find it more efficient to rely on readily observable characteristics (such as the status of the brokerage employing the analyst) as proxies for the accuracy of the forecast, rather than attempting to identify accurate analysts by monitoring analyst characteristics. In this case, analyst affiliation could be more strongly associated with return responses to forecast revisions than would be suggested by the effect of broker affiliation on predicted forecast accuracy alone.

Third, prior research shows that investors do not always weight information rationally in relation to its ability to predict future earnings. For example, prior studies show that investors appear to undervalue cash flows, and to overvalue accruals (Sloan 1996) and abnormal accruals (Xie 2001), relative to their ability to predict earnings. Similarly, Elgers et al. (2001) find that investors appear to undervalue the information in analysts' annual earnings forecasts for predicting future earnings.[2] Therefore, it is an empirical question whether the weights investors appear to place on forecasts diverge from the weights suggested by a model that predicts forecast accuracy.

This study extends prior research by conducting a more complete examination of the association between analyst accuracy and return responses to forecast revisions. Specifically, we examine whether investors respond to analysts' forecast revisions as if forecast accuracy is all that matters. We investigate the stock market's response to individual analysts' forecast revisions using a two-stage rational expectations procedure adapted from Mishkin (1983). In the first stage, we simultaneously estimate two models. First we estimate a forecast accuracy model to obtain a rational expectations prediction of analyst forecast accuracy. *Rational expectations* implies that investors use all available information to predict forecast accuracy, so our forecast accuracy model incorporates a broad set of explanatory variables identified in prior research, as well as a new variable that we expect to be associated with forecast accuracy (days elapsed since the prior forecast). We then estimate a valuation model that regresses the return response to a forecast revision on the magnitude of the forecast revision, weighted by the forecasting analyst's predicted forecast accuracy (from the forecast accuracy model). That is, we constrain the returns model coefficients to equal the forecast accuracy model coefficients. In the second stage, we estimate an unconstrained returns model that removes the forecast-accuracy-related constraints on estimated coefficients. We then determine whether the coefficients on analyst characteristics differ significantly across the constrained and unconstrained returns models. This procedure enables us to infer whether investors respond to forecast revisions as if they weight those revisions in proportion to the analyst's expected forecast accuracy.

---

[1] In our study, *timeliness* refers to how early or late the forecast date is in the fiscal period. Forecasts made earlier in the year are more timely and have longer horizons. The other predictors of forecast accuracy (firm-specific experience, broker size, forecast frequency, number of firms and industries followed) are analyst characteristics rather than forecast characteristics. For conciseness, we refer to the full set of potential explanatory variables, including forecast timeliness, as *analyst characteristics*.

[2] Our goal is to test whether price reactions to analyst forecasts suggest that investors weight forecasts based on rational predictions of forecast accuracy. We do not hypothesize or test market efficiency except in this narrow sense. Furthermore, evidence that investors do not focus solely on predicted forecast accuracy when responding to forecast revisions need not imply market inefficiency. For example, more timely but less accurate forecasts may provide useful information to investors. The consistency of return responses with market efficiency is an interesting issue for future research.

We find that the implied weights on the characteristics differ significantly across the constrained and unconstrained returns models. Specifically, security return responses are greater for timely (early) forecasts than for later forecasts, even though timely forecasts are generally less accurate than later forecasts. Sensitivity analysis incorporating a proxy for the level of uncertainty suggests that return responses to timely forecasts are stronger because there is greater uncertainty about earnings early in the year. Return responses to forecast revisions also increase with days elapsed since the prior forecast, even though these forecasts are, on average, less accurate than forecasts issued shortly after the prior forecast. In addition, forecasts provided by analysts employed by larger brokers are associated with greater return responses than would be predicted by those forecasters' accuracy alone. This result suggests that investors may use analyst affiliation as a signal of forecast accuracy, consistent with the findings of Hong and Kubik (2003). These results hold only for *high-innovation forecasts*—forecasts that are above both the analyst's prior forecast and the consensus forecast, or else are below both the analyst's prior forecast and the consensus (Gleason and Lee 2003). In contrast, for the remaining (*low-innovation*) forecast revisions, we find that the return response is not significantly related to any of the analyst characteristics that are predictive of forecast accuracy.

Our results make three contributions to the existing literature. First, we document that a broad set of analyst and forecast characteristics contributes simultaneously to explaining analyst forecast accuracy. Brown (2001) finds that past performance better predicts forecast accuracy than several analyst characteristics combined (general experience, company-specific experience, the number of companies and industries the analyst follows, and brokerage size). However, we find that a variable Brown (2001) omits—forecast frequency—and which Jacob et al. (1999) use to proxy the amount of effort the analyst devotes to following the company, is just as important as past accuracy for explaining future accuracy. Second, we show that forecast accuracy is not the sole determinant of security return responses to analyst forecast revisions. For example, returns respond more strongly to revisions by analysts employed by large brokerages than seems warranted by those analysts' forecast accuracy. On the other hand, forecast frequency, firm experience, and the number of companies and industries the analyst follows are associated with forecast accuracy, but not with the security price response to forecast revisions. More surprisingly, the forecast horizon and days elapsed since the last forecast are significantly negatively associated with forecast accuracy, but are positively associated with return responses, i.e., in a direction opposite to the forecast accuracy predictions. Third, we provide evidence on Schipper's (1991) questions about how investors trade off accuracy and timeliness; our results suggest that timeliness is important to investors even at the cost of accuracy.[3] Collectively, these results suggest that investors do *not* respond to analyst forecasts as if forecast accuracy is all that matters.

Our results also have implications for identifying proxies for expected earnings. The most common proxy, the mean forecast, implicitly assumes that investors weight all analysts' forecasts equally when forming expectations. However, our results suggest that investors apply different weights to individual forecasts. Researchers could use these weights to derive consensus forecasts that are closer to investors' expectations than the mean forecast. Managers could use the weights to more accurately identify investors' expectations in order to implement more effective disclosure strategies. For example, if investors rely heavily on forecasts of analysts employed by large brokers, then managers could compare their private information about upcoming earnings to those specific analysts' forecasts when deciding whether to make voluntary disclosures that may influence current expectations. Our results also have implications for evaluating analysts' performance. For example, the *Wall Street Journal* rates analysts based solely on forecast accuracy, but our results suggest that timeliness is also important to investors.

We organize the remainder of the paper as follows. Section II discusses prior literature that investigates the association between analysts' characteristics and forecast accuracy. Section III outlines the research design, Section IV discusses the results, and Section V summarizes our conclusions.

II. ANALYST-CHARACTERISTICS LITERATURE

Prior research documents that forecast frequency, broker size, general and firm-specific experience, and the number of companies and industries the analyst follows are associated with forecast accuracy. Jacob et al. (1999) find that forecast accuracy increases with forecast frequency (a proxy for the amount of effort the analyst devotes to following a company), even after controlling for forecast timeliness. Clement (1999) and Jacob et al. (1999) find that analysts employed by larger brokers issue more accurate forecasts on average. Analysts who issue more accurate forecasts are more likely to move to a high-status (i.e., large and prestigious) brokerage firm, and high-status brokers are more likely than other employers to dismiss an analyst for poor forecasting performance (Hong and Kubik 2003). Mikhail et al. (1997) and Jacob et al. (1999) find that forecast accuracy increases as analysts gain firm-specific experience, but they find no association for general forecasting experience. Clement (1999) finds that forecast accuracy increases with both firm-specific and general forecasting experience, although he finds that an additional year of firm-specific forecasting experience is associated with greater improvement in forecast accuracy than is an additional year of general forecasting experience. Jacob et al. (1999) and Clement (1999) also find that forecast accuracy declines as analysts follow more companies and industries.

Brown (2001) extends prior studies by investigating how well all of the abovementioned analyst characteristics—except forecast frequency—predict forecast accuracy. He finds that each characteristic provides incremental predictive power for forecast accuracy beyond past forecasting performance, although the analyst's past forecasting performance has greater predictive power than all of the other analyst characteristics combined. Thus, our forecast accuracy model includes characteristics in addition to past forecast accuracy because past accuracy is a noisy measure of forecasting ability. For example, able forecasters occasionally make a large error, and inept forecasters sometimes forecast accurately simply by chance.

---

[3] Analysts' role in setting stock prices may well have changed since the passage of Regulation Full Disclosure (FD). Prior to FD there is evidence that managers guided analysts (Matsumoto 2002), and so analysts' forecast revisions may have reflected managers' private information. If this guidance increased the value-relevance of analysts' forecast revisions, then analysts' forecast revisions may be less informative to investors in the post-FD environment, and our inferences may therefore no longer hold in that environment. Whether the informativeness of analysts' forecast revisions has changed with FD is an empirical question, and so far the results are mixed (Mohanram and Sunder 2002; Zitzewitz 2002; Agrawal and Chadra 2002; Heflin et al. 2003).

Although several studies investigate the association between analyst characteristics and forecast accuracy, we know of no prior evidence on whether investor responses to forecast revisions are consistent with rational estimates of analyst forecast accuracy. The following section outlines our hypotheses and research design for investigating this question.

III. SPECIFICATION OF EMPIRICAL TESTS

Analysts' Characteristics and Forecast Accuracy

To determine the association between analyst characteristics and forecast accuracy we begin with Brown's (2001) model of forecast accuracy as a function of firm-specific experience, broker size, prior forecast accuracy, forecast horizon, and the number of companies and industries the analyst follows. We add to this model forecast frequency (Jacob et al. 1999) and a new variable: days elapsed since the last forecast. Forecasts that are made long after the prior forecast are likely to reflect new information that was not incorporated in prior forecasts. We describe the variables in greater detail below.

To facilitate comparisons of estimated coefficients across different determinants of forecast accuracy within the forecast accuracy model, and between coefficients in the forecast accuracy and stock returns models, we standardize all analyst and forecast characteristics to lie between 0 and 1. The original minimum and maximum values for each explanatory variable correspond to transformed values of 0 and 1, respectively. For each analyst characteristic and analyst i, the standardized independent variables take the form:

$$\mathrm{CHARACTERISTIC}_{ijt} = \frac{\mathrm{RAW\_CHARACTERISTIC}_{ijt} - \mathrm{RAW\_CHARACTERISTICmin}_{jt}}{\mathrm{RAW\_CHARACTERISTICmax}_{jt} - \mathrm{RAW\_CHARACTERISTICmin}_{jt}},$$

where $\mathrm{RAW\_CHARACTERISTICmax}_{jt}$ and $\mathrm{RAW\_CHARACTERISTICmin}_{jt}$ are the original maximum and minimum values, respectively, of a characteristic for firm j in year t. A high value for $\mathrm{CHARACTERISTIC}_{ijt}$ indicates that analyst i scores high on that characteristic relative to other analysts who follow firm j in year t. Our standardization results in a relative measure for all analysts who followed the same firm during the same time period.[4] This approach controls for systematic firm-year differences in the characteristics.[5]

To obtain measures of forecasting performance that increase with forecast accuracy, we standardize the forecast accuracy measure to be 0 for the least accurate forecast (highest absolute forecast error) and 1 for the most accurate forecast (lowest absolute forecast error). The standardized accuracy measure for analyst i takes the form:

---

[4] Using dummy variables to specify whether an observation is above or below its mean retains less information about variation in the variables than the approach we use. Nonetheless, we obtain similar results using the dummy variable approach.

[5] As an example, suppose analyst i follows 20 companies and the minimum and maximum number of companies followed by analysts who follow firm j in year t are 5 and 25, respectively. Then analyst i falls at the 75 percent level [(20 − 5)/(25 − 5)] of the range of companies followed by analysts who followed firm j in year t.

$$\text{ACCURACY}_{ijt} = \frac{\text{AFEmax}_{jt} - \text{AFE}_{ijt}}{\text{AFEmax}_{jt} - \text{AFEmin}_{jt}},$$

where AFEmax_{jt} and AFEmin_{jt} are the maximum and minimum absolute forecast errors, respectively, for analysts following firm j in year t. Higher values for ACCURACY_{ijt} therefore indicate that the analyst is more accurate.

Analyst Characteristics and Stock Market Responses

If investors respond to an analyst's forecast revisions without considering the predicted accuracy of that analyst's forecast, then we can model investors' return response as:[6]

$$\text{CAR}_{ijt} = \delta_0 + \delta_1\text{REVP}_{ijt} + \varepsilon_{ijt}, \quad (1)$$

where CAR_{ijt} (cumulative abnormal return) is the three-trading-day cumulative market-adjusted abnormal return surrounding analyst i's earnings forecast revision for firm j in year t. The accumulation period starts one day before the revision and ends one day after the revision. We compute the return by subtracting value-weighted market returns from firm j's return. REVP_{ijt} (revision scaled by price) is analyst i's forecast revision for firm j in year t scaled by the end-of-day stock price two days prior to the revision. The revision equals analyst i's forecast of firm j's earnings in year t, less analyst i's prior forecast of firm j's earnings in year t.[7]

If, however, investors respond to forecast revisions based on the analyst's predicted accuracy, then we can model investor response to forecast revisions as:

$$\text{CAR}_{ijt} = \gamma_0 + \gamma_1\text{REVP}_{ijt} \times \text{ACCURACY}^{P}_{ijt} + \varepsilon_{ijt}, \quad (2)$$

where $\text{ACCURACY}^{P}_{ijt}$ = the predicted accuracy of the forecast revision.

Since future accuracy is unknown when the analyst issues a forecast, investors must use information available at the time of the forecast to estimate its accuracy. We predict that the analyst's forecasting performance improves with prior forecasting performance, broker size, forecast frequency, and experience forecasting for that firm, and decreases with days elapsed since the last forecast, forecast horizon, and the number of companies and industries the analyst follows:

$$\begin{aligned}\text{ACCURACY}_{ijt} = {} & \alpha_0 + \alpha_1\text{DAYS\_ELAPSED}_{ijt} + \alpha_2\text{FOR\_HORIZON}_{ijt} \\ & + \alpha_3\text{LAG\_ACCURACY}_{ijt} + \alpha_4\text{BROKER\_SIZE}_{ijt} \\ & + \alpha_5\text{FOR\_FREQUENCY}_{ijt} + \alpha_6\text{FIRM\_EXPERIENCE}_{ijt} \\ & + \alpha_7\text{COMPANIES}_{ijt} + \alpha_8\text{INDUSTRIES}_{ijt} + \varepsilon_{ijt}, \quad (3)\end{aligned}$$

---

[6] Because analysts issue forecasts throughout the year, it is unlikely that returns around forecast revisions are systematically related to other events. In particular, the returns are not likely to be due to management forecasts, which occur infrequently. Using the First Call database, we find 7,581 management forecasts of earnings for 2,899 firms from January 1993 to March 1998, an average of one management forecast every two years for each firm. In contrast, we find 265,620 analyst forecasts for 11,238 firms on the I/B/E/S database for about the same period, even after we restrict the sample to one forecast per analyst per year.

[7] We measure forecast revisions relative to the analyst's own prior forecast because that measure is more strongly correlated with returns than are revisions measured relative to the consensus forecast. Gleason and Lee (2003) find similar results, and recommend measuring forecast revisions relative to the analyst's own prior forecast. However, measuring the revision relative to the consensus forecast leads to identical inferences.

where:

ACCURACY_{ijt} = a measure of analyst i's forecast accuracy for firm j in year t. It is calculated as the maximum absolute forecast error for analysts who follow firm j in year t minus the absolute forecast error of analyst i following firm j in year t, with this difference scaled by the range of absolute forecast errors for analysts following firm j in year t;

$\text{DAYS\_ELAPSED}_{ijt}$ = a measure of the days elapsed since the last forecast by any analyst following firm j in year t. It is calculated as the days between analyst i's forecast of firm j's earnings in year t and the most recent preceding forecast of firm j's earnings by any analyst, minus the minimum number of days between two adjacent forecasts of firm j's earnings by any two analysts in year t, with this difference scaled by the range of days between two adjacent forecasts of firm j's earnings in year t;

FOR_HORIZON_{ijt} = a measure of the time from the forecast date to the end of the fiscal period. It is calculated as the forecast horizon (days from the forecast date to the fiscal year-end) for analyst i following firm j in year t minus the minimum forecast horizon for analysts who follow firm j in year t, with this difference scaled by the range of forecast horizons for analysts following firm j in year t;

$\text{LAG\_ACCURACY}_{ijt}$ = a measure of analyst i's prior forecast accuracy for firm j. It is calculated as the maximum ACCURACY for analysts who follow firm j in year $t - 1$ minus the ACCURACY for analyst i following firm j in year $t - 1$, with this difference scaled by the range of ACCURACY for analysts following firm j in year $t - 1$;

BROKER_SIZE_{ijt} = a measure of the analyst's broker size. It is calculated as the number of analysts employed by the broker employing analyst i following firm j in year t minus the minimum number of analysts employed by brokers for analysts following firm j in year t, with this difference scaled by the range of brokerage size for analysts following firm j in year t;

$\text{FOR\_FREQUENCY}_{ijt}$ = a measure of analyst i's forecast frequency for firm j. It is calculated as the number of firm-j forecasts made by analyst i following firm j in year t minus the minimum number of firm-j forecasts for analysts following firm j in year t, with this difference scaled by the range of number of firm-j forecasts issued by analysts following firm j in year t;

$\text{FIRM\_EXPERIENCE}_{ijt}$ = a measure of analyst i's firm-specific experience. It is calculated as the number of years of firm-specific experience for analyst i following firm j in year t minus the minimum number of years of firm-specific experience for analysts following firm j in year t, with this difference scaled by the range of years of firm-specific experience for analysts following firm j in year t;

$COMPANIES_{ijt}$ = a measure of the number of companies analyst i follows in year t. It is calculated as the number of companies followed by analyst i following firm j in year t minus the minimum number of companies followed by analysts who follow firm j in year t, with this difference scaled by the range in the number of companies followed by analysts following firm j in year t; and

$INDUSTRIES_{ijt}$ = a measure of the number of industries analyst i follows in year t. It is calculated as the number of two-digit SICs followed by analyst i following firm j in year t minus the minimum number of two-digit SICs followed by analysts who follow firm j in year t, with this difference scaled by the range in the number of two-digit SICs followed by analysts following firm j in year t.

We refer to Equation (3) as the *forecast accuracy model.*

Substituting the predicted forecast accuracy from Equation (3) into Equation (2) yields the returns model:

$$\begin{aligned} CAR_{ijt} = \gamma_0 + \gamma_1 REVP_{ijt} \times (\alpha_0^* &+ \alpha_1^* DAYS\_ELAPSED_{ijt} + \alpha_2^* FOR\_HORIZON_{ijt} \\ &+ \alpha_3^* LAG\_ACCURACY_{ijt} + \alpha_4^* BROKER\_SIZE_{ijt} \\ &+ \alpha_5^* FOR\_FREQUENCY_{ijt} + \alpha_6^* FIRM\_EXPERIENCE_{ijt} \\ &+ \alpha_7^* COMPANIES_{ijt} + \alpha_8^* INDUSTRIES_{ijt}) + \varepsilon_{ijt}, \end{aligned} \tag{4}$$

where the α_i^* are the returns model coefficients corresponding to the forecast model coefficients (α_i) for each analyst characteristic in Equation (3). We refer to Equation (4) as the *valuation model.*

Using the method outlined in Mishkin (1983), we estimate the forecast accuracy and valuation models in two stages using iterative weighted nonlinear least squares.[8] In the first stage of the model we constrain $\alpha_i^* = \alpha_i$. In the second stage we reestimate Equation (4) and allow the coefficients to vary freely.[9] If investor responses to forecast revisions reflect the ability of analyst characteristics to predict forecast accuracy, then α_i^* should equal α_i

---

[8] Sloan (1996) introduced the Mishkin (1983) model to accounting research in his study investigating whether investor responses to earnings are consistent with rational expectations, based on the ability of cash flow and accruals to predict next-period income. Xie (2001) and Elgers et al. (2001) also use the Mishkin model to study the consistency of returns with rational expectations of earnings from abnormal accruals and analyst forecasts, respectively. We also estimated the forecast accuracy model and the constrained and unconstrained returns models (Equations [3] and [4]) using ordinary least squares (OLS), and tested the constraint on the returns model coefficients. Both the OLS and Mishkin approaches yield qualitatively similar results. We conclude that the study's results are not induced by the Mishkin approach.

[9] This second stage can be written as the OLS model:

$$\begin{aligned} CAR_{ijt} = \delta_0 + \beta_0 REVP_{ijt} &+ \beta_1 REVP_{ijt} \times DAYS\_ELAPSED_{ijt} \\ &+ \beta_2 REVP \times FOR\_HORIZON_{ijt} + \beta_3 REVP_{ijt} \times LAG\_ACCURACY_{ijt} \\ &+ \beta_4 REVP_{ijt} \times BROKER\_SIZE_{ijt} + \beta_5 REVP_{ijt} \times FOR\_FREQUENCY_{ijt} \\ &+ \beta_6 REVP_{ijt} \times FIRM\_EXPERIENCE_{ijt} + \beta_7 REVP_{ijt} \times COMPANIES_{ijt} \\ &+ \beta_8 REVP_{ijt} \times INDUSTRIES_{ijt} + \varepsilon_{ijt}, \end{aligned}$$

where $\alpha_i^* = \beta_i/\gamma_1$, but we derive our results by jointly estimating Equations (3) and (4) using the Mishkin procedure.

for all characteristics in the unconstrained returns model. In contrast, if the weights investors place on analyst characteristics when responding to forecast revisions differ from the characteristics' association with forecast accuracy, then the coefficients will differ between the constrained and unconstrained returns models.[10]

We assess the difference between coefficients in the constrained and unconstrained returns models using the following likelihood-ratio statistic, which is distributed asymptotically as $\chi^2(q)$: $-2n \log (SSR^c/SSR^u)$, where q = the number of constraints; n = the number of observations; SSR^c = the sum of squared residuals from the constrained estimation; and SSR^u = the sum of squared residuals from the unconstrained estimation.

IV. SAMPLE SELECTION AND RESULTS

Sample Selection

Discussions with I/B/E/S officials indicated that prior to the early 1990s, the forecast release date I/B/E/S recorded often differed from the actual forecast date by a few days, but that the accuracy of the forecast dates improved in the early 1990s. Clearly, we need accurate release dates to measure return responses to the forecast revisions. Thus, our sample includes annual earnings forecasts from the I/B/E/S data set for the period 1994–1998.

We restrict the sample to forecasts made at least 30 days, but not more than one year, before the end of the accounting period (so the minimum forecast horizon is 30 days). Consistent with other studies (e.g., O'Brien 1990; Sinha et al. 1997; Clement 1999), we retain only the last forecast for each analyst-firm pair during the period.[11] We obtain returns data from CRSP. We eliminate observations from the sample if (1) there are no CRSP returns data around the forecast revision, (2) the analyst did not issue a forecast in the prior year for the firm (the analyst's prior forecast performance is an independent variable in the model), (3) the scaled forecast revision is in the top or bottom 1 percent of revisions, or (4) only one analyst follows the firm. The last requirement is necessary because we measure an analyst's characteristics and forecast accuracy relative to other analysts who follow the same firm.[12]

Results

Descriptive Statistics

Table 1 reports summary statistics for the sample. Panel A shows the distributions of the raw (untransformed) variables. The distributions are consistent with those reported for similar variables in other studies (Clement 1999; Jacob et al. 1999). For example, forecasts precede the fiscal year-end by 97 days, on average (recall that our minimum forecast horizon is 30 days). Each analyst provides about four forecasts for each firm in a given year, and analysts follow an average of 20 firms in five industries. The median values of the firm characteristics are generally different from the means, indicating that the distributions of

---

[10] The test described is valid even if the forecasting equation is incomplete, as Mishkin (1983, 45–46) explains: "Thus leaving out relevant variables from the OLS regression...will not affect the rationality implication that $(\alpha_1 - \alpha_1^*)$ should not differ significantly from zero." He goes on to state that the above procedure tests rationality no matter what available past information is included in or excluded from the forecasting equation. Sloan (1996, 302, footnote 13) makes a similar point.

[11] Because our objective is to determine whether investors respond to characteristics that are associated with systematic differences in forecast accuracy, we need only one analyst-firm pair for each analyst-firm-year. Prior studies (e.g., Clement 1999; Brown 2001) suggest that this sample selection method identifies samples with systematic differences in forecast accuracy.

[12] Restricting the sample to firms with forecasts from a minimum of three or four analysts per firm instead of two does not affect our inferences.

TABLE 1
Descriptive Statistics on Analyst Characteristics for Sample Firms, 1994–1998
(n = 35,351)

Panel A: Distribution of Selected Analyst Characteristics before Standardization[a]

Variable	*Units of Measurement*	*Mean*	*25th Percentile*	*Median*	*75th Percentile*
DAYS_ELAPSED	Days from prior forecast	13.6	2	6	16
FOR_HORIZON	Days to fiscal year-end	97.0	57	71	131
BROKER_SIZE	Number of analysts	33.8	12	29	49
FOR_FREQUENCY	Number of forecasts	3.8	2	3	5
FIRM_EXPERIENCE	Years	3.9	2	3	5
COMPANIES	Companies followed	19.5	11	16	23
INDUSTRIES	Industries followed	5.4	3	5	7

Panel B: Distribution of Analyst Characteristics after Standardization

Variable	*Mean*	*25th Percentile*	*Median*	*75th Percentile*
ACCURACY	0.58	0.06	0.71	1.00
DAYS_ELAPSED	0.38	0.00	0.20	0.82
FOR_HORIZON	0.38	0.03	0.20	0.82
LAG_ACCURACY	0.57	0.00	0.67	1.00
BROKER_SIZE	0.43	0.05	0.35	0.82
FOR_FREQUENCY	0.42	0.00	0.33	0.80
FIRM_EXPERIENCE	0.36	0.00	0.20	0.71
COMPANIES	0.41	0.05	0.30	0.75
INDUSTRIES	0.40	0.00	0.29	0.75

(continued on next page)

TABLE 1 (continued)

Panel C: Pearson Correlations among Analyst Characteristics (Significance levels are in parentheses)

	ACCURACY	*DAYS_ ELAPSED*	*FOR_ HORIZON*	*LAG_ ACCURACY*	*BROKER_ SIZE*	*FOR_ FREQUENCY*	*FIRM_ EXPERIENCE*	*COMPANIES*
DAYS_ELAPSED	−0.047 (<0.001)							
FOR_HORIZON	−0.251 (<0.001)	0.055 (<0.001)						
LAG_ACCURACY	0.084 (<0.001)	−0.034 (<0.001)	−0.042 (<0.001)					
BROKER_SIZE	0.019 (<0.001)	0.036 (<0.001)	0.032 (<0.001)	0.020 (<0.001)				
FOR_FREQUENCY	0.146 (<0.001)	0.006 (0.260)	−0.339 (<0.001)	0.033 (<0.001)	0.017 (<0.001)			
FIRM_EXPERIENCE	0.019 (<0.001)	0.001 (0.822)	−0.018 (<0.001)	0.003 (0.607)	−0.002 (0.755)	0.018 (<0.001)		
COMPANIES	−0.036 (<0.001)	0.023 (<0.001)	−0.017 (0.002)	−0.055 (<0.001)	0.007 (0.159)	−0.008 (0.139)	0.154 (<0.001)	
INDUSTRIES	−0.040 (<0.001)	0.010 (0.065)	−0.000 (0.947)	−0.050 (<0.001)	−0.118 (<0.001)	−0.035 (<0.001)	0.091 (<0.001)	0.544 (<0.001)

(continued on next page)

TABLE 1 (continued)

---

[a] We omit forecast accuracy from Panel A because the unstandardized variable (absolute per-share forecast error) is not directly comparable among firms. We report descriptive statistics for standardized forecast errors in Panel B.

ACCURACY_{ijt} = a measure of analyst i's forecast accuracy for firm j in year t. It is calculated as the maximum absolute forecast error for analysts who follow firm j in year t minus the absolute forecast error of analyst i following firm j in year t, with this difference scaled by the range of absolute forecast errors for analysts following firm j in year t;

$\text{DAYS\_ELAPSED}_{ijt}$ = a measure of the days elapsed since the last forecast by any analyst following firm j in year t. It is calculated as the days between analyst i's forecast of firm j's earnings in year t and the most recent preceding forecast of firm j's earnings by any analyst, minus the minimum number of days between two adjacent forecasts of firm j's earnings by any two analysts in year t, with this difference scaled by the range of days between two adjacent forecasts of firm j's earnings in year t;

FOR_HORIZON_{ijt} = a measure of the time from the forecast date to the end of the fiscal period. It is calculated as the forecast horizon (days from the forecast date to the fiscal year-end) for analyst i following firm j in year t minus the minimum forecast horizon for analysts who follow firm j in year t, with this difference scaled by the range of forecast horizons for analysts following firm j in year t;

$\text{LAG\_ACCURACY}_{ijt}$ = a measure of analyst i's prior forecast accuracy for firm j. It is calculated as the maximum ACCURACY for analysts who follow firm j in year t − 1 minus the ACCURACY for analyst i following firm j in year t − 1, with this difference scaled by the range of ACCURACY for analysts following firm j in year t − 1;

BROKER_SIZE_{ijt} = a measure of the analyst's broker size. It is calculated as the number of analysts employed by the broker employing analyst i following firm j in year t minus the minimum number of analysts employed by brokers for analysts following firm j in year t, with this difference scaled by the range of brokerage size for analysts following firm j in year t;

$\text{FOR\_FREQUENCY}_{ijt}$ = a measure of analyst i's forecast frequency for firm j. It is calculated as the number of firm-j forecasts made by analyst i following firm j in year t minus the minimum number of firm-j forecasts for analysts following firm j in year t, with this difference scaled by the range of number of firm-j forecasts issued by analysts following firm j in year t;

$\text{FIRM\_EXPERIENCE}_{ijt}$ = a measure of analyst i's firm-specific experience. It is calculated as the number of years of firm-specific experience for analyst i following firm j in year t minus the minimum number of years of firm-specific experience for analysts following firm j in year t, with this difference scaled by the range of years of firm-specific experience for analysts following firm j in year t;

COMPANIES_{ijt} = a measure of the number of companies analyst i follows in year t. It is calculated as the number of companies followed by analyst i following firm j in year t minus the minimum number of companies followed by analysts who follow firm j in year t, with this difference scaled by the range in the number of companies followed by analysts following firm j in year t; and

INDUSTRIES_{ijt} = a measure of the number of industries analyst i follows in year t. It is calculated as the number of two-digit SICs followed by analyst i following firm j in year t minus the minimum number of two-digit SICs followed by analysts who follow firm j in year t, with this difference scaled by the range in the number of two-digit SICs followed by analysts following firm j in year t.

the characteristics are asymmetric. Panel B shows the distributions of the variables in Panel A, standardized to lie between 0 and 1. This transformation preserves the relative values of the explanatory variables while reducing the influence of extreme observations.[13]

Panel C of Table 1 reports correlations among the variables. All of the correlations between the analyst characteristics and forecast accuracy are significant in the expected directions at better than the 1 percent level. The characteristics that are associated with greater forecast accuracy are (from highest to lowest correlation) forecast frequency, lagged forecast accuracy, broker size, and firm-specific experience. The characteristics that are associated with lower forecast accuracy are (from the most negative to the least negative correlation) forecast horizon, days elapsed since the last forecast, number of industries the analyst follows, and number of companies the analyst follows. The number of companies and industries the analyst follows are highly positively correlated, while forecast frequency is negatively correlated with forecast horizon. The positive correlation between firm experience and companies followed suggests that more experienced analysts follow more firms. The negative correlation between broker size and industries followed suggests that larger brokerages permit greater specialization among their analyst employees.

Tests of Association between Analysts' Characteristics and Forecast Accuracy

Panel A of Table 2 reports the results of estimating the forecast accuracy model (Equation [3]). Consistent with the correlations in Table 1, the regression results show that an analyst's accuracy (relative to that of other analysts forecasting the same firm's earnings in the same year) increases with prior accuracy, broker size, forecast frequency, and firm experience, and decreases with days elapsed since the last forecast, forecast horizon, and the number of companies and industries the analyst follows. These results are generally consistent with prior studies, indicating that our transformation of the variables to lie between 0 and 1 preserves the information in analyst forecasts and characteristics. However, our results yield two new insights. First, whereas Brown (2001) finds that past forecast accuracy dominates individual analyst characteristics in explaining current forecast accuracy, we find that a variable Brown (2001) omits—the frequency with which the analyst forecasts the firm's earnings, which Jacob et al. (1999) use to proxy the amount of effort the analyst devotes to following the company—is just as important as past forecast accuracy for explaining current forecast accuracy. The forecast model coefficients for past forecast accuracy and forecast frequency are 0.068 and 0.067, respectively, and the difference between them is statistically insignificant. Second, we find that forecast accuracy decreases with days elapsed since the last forecast—a new forecast characteristic not investigated in prior research. In sum, our results suggest that investors can use these characteristics of analysts and forecasts to predict forecast accuracy.

Tests of Association between Analysts' Characteristics and Stock Returns

To investigate whether investors condition their responses to forecast revisions in a manner consistent with the analyst's expected accuracy, we compare the coefficients from the forecast accuracy model (Equation [3]), which by construction has the same coefficients as the constrained returns model, with the corresponding coefficients from the unconstrained

---

[13] A median standardized value of 0.5 would indicate that half of the firm-year observations are above the midpoint of the range for that variable. The medians in Table 1, Panel B show that for some variables, half of the observations are in the bottom 20 percent of the range of the variables. Our forecast accuracy regression model does not require the explanatory variables to conform to any particular distribution. However, our hypothesis tests assume that the distribution of the regression error term (i.e., the distribution of the dependent variable, forecast accuracy, conditional on the distribution of the explanatory variables) is normal.

TABLE 2
Regression Model Coefficient Estimates for the Forecast Accuracy and Returns Models
(all forecast revisions: n = 35,351)

$$\begin{aligned} ACCURACY_{ijt} = {} & \alpha_0 + \alpha_1 DAYS\_ELAPSED_{ijt} + \alpha_2 FOR\_HORIZON_{ijt} \\ & + \alpha_3 LAG\_ACCURACY_{ijt} + \alpha_4 BROKER\_SIZE_{ijt} \\ & + \alpha_5 FOR\_FREQUENCY_{ijt} + \alpha_6 FIRM\_EXPERIENCE_{ijt} \\ & + \alpha_7 COMPANIES_{ijt} + \alpha_8 INDUSTRIES_{ijt} + \varepsilon_{ijt} \end{aligned} \quad (3)$$

$$\begin{aligned} CAR_{ijt} = {} & \gamma_0 + \gamma_1 REVP_{ijt} \times (\alpha_0^* + \alpha_1^* DAYS\_ELAPSED_{ijt} + \alpha_2^* FOR\_HORIZON_{ijt} \\ & + \alpha_3^* LAG\_ACCURACY_{ijt} + \alpha_4^* BROKER\_SIZE_{ijt} \\ & + \alpha_5^* FOR\_FREQUENCY_{ijt} + \alpha_6^* FIRM\_EXPERIENCE_{ijt} \\ & + \alpha_7^* COMPANIES_{ijt} + \alpha_8^* INDUSTRIES_{ijt}) + \varepsilon_{ijt} \end{aligned} \quad (4)$$

Panel A: Parameter Estimates

	Forecast Model Coefficients			*Returns Model Coefficients*		
	Parameter	*Estimate*	*Asymptotic Std. Error*	*Parameter*	*Estimate*	*Asymptotic Std. Error*
DAYS_ELAPSED	α_1	−0.032***	0.0051	α_1^*	0.236***	0.0567
FOR_HORIZON	α_2	−0.229***	0.0055	α_2^*	0.296***	0.0604
LAG_ACCURACY	α_3	0.068***	0.0051	α_3^*	0.101*	0.0571
BROKER_SIZE	α_4	0.024***	0.0054	α_4^*	0.391***	0.0601
FOR_FREQUENCY	α_5	0.067***	0.0055	α_5^*	0.084	0.0627
FIRM_EXPERIENCE	α_6	0.020***	0.0052	α_6^*	0.048	0.0575
COMPANIES	α_7	−0.030***	0.0066	α_7^*	−0.000	0.0728
INDUSTRIES	α_8	−0.019***	0.0064	α_8^*	−0.045	0.0711
REVP[a]				γ_1	3.123***	0.1230

Panel B: Tests of Equality of Forecast and Returns Model Coefficients

	Parameters	*Likelihood Ratio Statistic*	*Marginal Significance Level*
DAYS_ELAPSED	$\alpha_1 = \alpha_1^*$	8.53	0.004
FOR_HORIZON	$\alpha_2 = \alpha_2^*$	21.77	<0.001
LAG_ACCURACY	$\alpha_3 = \alpha_3^*$	1.02	0.313
BROKER_SIZE	$\alpha_4 = \alpha_4^*$	18.93	<0.001
FOR_FREQUENCY	$\alpha_5 = \alpha_5^*$	0.81	0.369
FIRM_EXPERIENCE	$\alpha_6 = \alpha_6^*$	0.66	0.418
COMPANIES	$\alpha_7 = \alpha_7^*$	1.86	0.174
INDUSTRIES	$\alpha_8 = \alpha_8^*$	0.44	0.491
Overall (joint) test	$\alpha_i = \alpha_i^*$, i = 1,...,8	67.99	<0.001

(continued on next page)

TABLE 2 (continued)

---

*, *** Significantly different from zero at the 10, 5, and 1 percent levels, respectively.

[a] We estimate γ_i when we constrain the α_i^* coefficient in Equation (4) to equal the corresponding α_i coefficients from Equation (3).

CAR_{ijt} (cumulative abnormal return) = three-trading-day cumulative market-adjusted abnormal return surrounding analyst i's earnings forecast revision for firm j in year t. The accumulation period starts one day before the revision and ends one day after the revision. We compute the return by subtracting value-weighted market returns from firm j's return; and

$REVP_{ijt}$ (revision scaled by price) = analyst i's forecast revision for firm j in year t scaled by the end-of-day stock price two days prior to the revision. The revision equals analyst i's forecast of firm j's earnings in year t less analyst i's prior forecast of firm j's earnings in year t.

Other variables are defined in Table 1.

returns model (Equation [4]), also reported in Table 2, Panel A.[14] If investors weight forecast revisions in accordance with rational expectations of forecast accuracy (as modeled in Equation [3]), then the coefficients for each of the characteristics should be the same in the forecast accuracy model and the unconstrained returns model (Equations [3] and [4], respectively).

All of the characteristics have significant coefficients in the forecast accuracy model, but only four characteristics have statistically significant coefficients in the unconstrained returns model: days elapsed, forecast horizon, past accuracy, and broker size. All four are positively associated with return responses to forecast revisions. Note, however, that both days elapsed and forecast horizon are associated with *less* accurate forecasts in the forecast accuracy model.

The remaining analyst characteristics are statistically insignificant in the unconstrained returns model, indicating that investors act as if they condition their responses to forecast revisions on only a subset of the characteristics that predict the revision's accuracy. One explanation for this divergence between the forecast accuracy model and unconstrained returns model estimates is that investors believe that the cost of monitoring these analyst-specific details exceeds the benefit from improved estimates of forecast accuracy that monitoring may provide.[15]

Panel B of Table 2 formally tests differences in coefficients estimated in the constrained and unconstrained returns models by comparing sums of squares for restricted and unrestricted returns models. After estimating the unrestricted model (which allows the returns model coefficients to differ from their forecast accuracy model counterparts), we then compute the restricted returns model by restricting the coefficient of each characteristic in turn to be equal in the forecast accuracy and stock returns models. We compute a Chi-square statistic based on the increase in residual sums of squares resulting from the restriction. The individual coefficients for days elapsed, forecast horizon, and broker size differ significantly in the two models at better than the 1 percent level. The positive coefficient for forecast horizon in the unconstrained returns model suggests that investors respond strongly to timely forecasts, even though the forecast accuracy model reveals that those longer-horizon forecasts are generally less accurate than later forecasts. Investors also respond strongly to forecasts that break long absences of analyst forecasts for a given firm, even though those forecasts are also less accurate than other forecasts. In addition, investors appear to place more weight on broker size than seems warranted by the forecast accuracy model's coefficients, consistent with investors' using analyst affiliation as a signal of forecast accuracy. The differences in the coefficients on the remaining characteristics are statistically insignificant.

---

[14] We obtain similar coefficients and significance levels when we estimate the forecast accuracy and returns models using OLS. The R^2s for the forecast accuracy and returns models are 0.076 and 0.022, respectively, which are consistent with prior forecast accuracy research (e.g., Clement 1999; Mikhail et al. 1997).

[15] Multicollinearity does not explain these variables' lack of significance in the stock returns model. We compute a Chi-square statistic to test the joint significance of the coefficients of forecast frequency, firm experience, and the number of companies and industries the analyst follows, in the stock returns model (α_5^* to α_8^*). The Chi-square statistic is 3.58, which is not statistically significant (4 degrees of freedom). In contrast, the corresponding variables in the forecast accuracy model are highly significant, both individually and jointly.

We also test the overall (joint) restriction of all coefficients by restricting all of the characteristics' coefficients to be equal in the forecast and unconstrained returns models. This overall test shows that the forecast and returns model coefficients differ significantly at better than the 1 percent level. We conclude that return responses to forecast revisions are inconsistent with investors' use of rational expectations of forecast accuracy (as modeled in Equation [3]) to weight the revisions.[16]

Lagged forecast accuracy, a summary measure of forecast accuracy as documented in Brown (2001), and broker size, which investors may use as a readily observable signal of forecast accuracy (Hong and Kubik 2003) are the only characteristics that are significantly associated with both forecast accuracy and stock returns in a consistent manner (i.e., with increased forecast accuracy *and* increased return responses). We therefore investigate whether the relative importance of these two key analyst characteristics differs in the forecast accuracy and unconstrained returns models. In the forecast accuracy model, the lagged forecast accuracy coefficient is significantly *higher* than the broker-size coefficient, but in the returns model, the lagged forecast accuracy coefficient is significantly *lower* than the broker-size coefficient (both at better than the 1 percent significance level).[17] These results reinforce our conclusion that investors do not respond solely to prior or predicted forecast accuracy. Furthermore, investors appear to respond more strongly to an easily observable analyst characteristic (broker size) than to a less easily observed characteristic (lagged forecast accuracy), even though lagged forecast accuracy is more strongly associated with forecast accuracy. This result suggests that investors use broker size as a summary indicator of analyst accuracy, and is consistent with Hong and Kubik's (2003) conclusion that large brokerage firms tend to hire and retain accurate forecasters.

In summary, the results in Table 2 suggest that forecast frequency, firm experience, and the number of companies and industries the analyst follows are significantly associated with forecast accuracy but not with returns. Although lagged forecast accuracy is more important than broker size in explaining forecast accuracy, it is less important than broker size in explaining the magnitude of the price reactions to forecast revisions. More surprisingly, forecast horizon and days elapsed since the last forecast are both associated with less accurate forecasts, but with larger price reactions to forecast revisions. Collectively, these results support our conclusion that investor responses to analysts' forecast revisions reflect characteristics other than predicted forecast accuracy. Specifically, investors respond strongly to timely forecasts and appear to use broker size as a proxy for forecast accuracy.

---

[16] We base this conclusion on the overall test in Panel B of Table 2, which shows that the forecast and returns model coefficients are significantly different. Although some individual characteristics (such as past accuracy, forecast frequency, and firm experience) have statistically similar coefficients in the forecast accuracy and returns models, the results in Panel A show that forecast frequency, firm experience, and companies and industries followed are not significantly associated with returns, indicating that investors respond to just a subset of the characteristics that are associated with forecast accuracy. Predictions of forecast accuracy that are based on only a subset of informative analyst characteristics are inconsistent with rational expectations because they ignore relevant information.

[17] We can compare the coefficient estimates because we scale both variables to range from 0 to 1. The coefficient for each scaled variable measures the return response to forecast revisions when that variable is at its highest level within a firm-year (i.e., when the scaled explanatory variable is 1).

Sensitivity Analysis

We perform several sensitivity analyses to check the robustness of the results. First, we examine whether controlling for (a proxy for) the level of prior uncertainty affects our inferences. Second, we partition the sample into high-innovation and low-innovation forecasts, as defined by Gleason and Lee (2003), to determine whether this partition drives our findings that returns respond more strongly to timely forecast revisions and revisions by analysts employed by large brokers. Finally, we estimate single-characteristic regressions as an additional test for possible effects of multicollinearity in the stock returns model. As explained below, our primary inferences remain unchanged after performing these tests.

Theory predicts that investor responses to forecast revisions increase with both the precision of the forecast and the precision of investors' prior information (Kim and Verrecchia 1991; Abarbanell et al. 1995). Our returns model focuses on the role of predicted forecast precision. To ensure that our results are not attributable to the omission of the precision of investors' prior information, or prior uncertainty, we reestimate the returns model after including a proxy for prior uncertainty. Following Barron et al.'s (1998) theoretical analysis of financial analysts' information environment, we measure prior uncertainty as the mean squared forecast error of all prior forecasts outstanding at each forecast date. Our uncertainty proxy measures the overall error among analysts, but is based on actual earnings information that is not available at the time of the forecast.[18] Similar to our other explanatory variables, we standardize prior uncertainty to range from 0 to 1 for each firm-year. We estimate the following model:

$$\begin{aligned} CAR_{ijt} = {} & \gamma_0 + \gamma_1 REVP_{ijt} \times (\alpha_0^* + \alpha_1^* DAYS\_ELAPSED_{ijt} + \alpha_2^* FOR\_HORIZON_{ijt} \\ & + \alpha_3^* LAG\_ACCURACY_{ijt} + \alpha_4^* BROKER\_SIZE_{ijt} \\ & + \alpha_5^* FOR\_FREQUENCY_{ijt} + \alpha_6^* FIRM\_EXPERIENCE_{ijt} \\ & + \alpha_7^* COMPANIES_{ijt} + \alpha_8^* INDUSTRIES_{ijt}) + \gamma_2 REVP_{ijt} \\ & \times UNCERTAINTY_{ijt} + \varepsilon_{ijt}, \end{aligned} \tag{5}$$

where:

$UNCERTAINTY_{ijt}$ = mean squared forecast error of analyst forecasts immediately prior to analyst i's forecast revision for firm j in year t, minus the minimum mean squared forecast error of forecasts immediately preceding the forecast revisions of analysts following firm j in year t, all scaled by the range of mean squared forecast errors preceding forecasts of analysts following firm j in year t.

We use the same prediction model for forecast accuracy as in the previous analysis.

---

[18] Barron et al. (1998, 424) label their uncertainty measure V, and define it as the *expected* squared forecast error from the earnings random variable. We use *realized* forecast errors because of the practical difficulty of measuring expected forecast errors for successive analyst forecasts. A possible alternative is to estimate uncertainty at a given date as the squared forecast error for forecasts outstanding on the same date in a preceding year. The disadvantage of that approach is that analysts' uncertainty may change from year to year with the firm's economic conditions.

However, the modified returns model in Equation (5) includes a response to predicted forecast accuracy, or forecast precision (measured with γ_1) and a response to prior uncertainty (measured with γ_2). In unreported results, we find that the coefficient on prior uncertainty (γ_2) is significantly positive. Interestingly, the forecast horizon coefficient is insignificant when we include prior uncertainty in the model.[19] Therefore, return responses to forecast revisions (1) increase with prior uncertainty, and (2) are not significantly associated with forecast horizon when we include prior uncertainty in the model. The coefficients on our other variables are essentially unchanged in this model, and our primary inferences are unaltered. Specifically, investor responses to forecast revisions are inconsistent with the use of analyst and forecast characteristics to derive rational expectations of analyst forecast accuracy.[20]

In our second sensitivity test we partition the sample in a manner similar to Gleason and Lee (2003). They find that high-innovation forecast revisions (new forecasts that are above *both* the analysts' prior forecast *and* the consensus forecast, or else are below *both* forecasts) have a greater price effect than low-innovation forecasts. This finding could explain our results if analyst characteristics such as forecast horizon and broker size are correlated with innovation. Thus, we estimate separately the forecast accuracy and stock returns models for high-innovation and low-innovation forecasts. We report the results in Tables 3 and 4, respectively.

Table 3, Panel A shows that the results for high-innovation forecasts are similar to the results for the full sample. Specifically, forecast frequency, firm experience, and the number of companies and industries the analyst follows are significantly associated with forecast accuracy but not with returns. Days elapsed since the prior forecast and forecast horizon are both associated with less accurate forecasts, but with larger price reactions to forecast revisions. Broker size is significantly associated with greater forecast accuracy and with larger price reactions to forecast revisions. This is the only analyst characteristic whose association with forecast accuracy is consistent with its association with returns. The results in Table 3, Panel B are also similar to the full sample results. The overall test shows that the coefficient estimates for analyst characteristics are significantly different in the forecast accuracy and returns models. The estimated coefficients for days elapsed, forecast horizon, broker size, and the number of companies followed are all significantly different in the forecast accuracy and returns models. Thus even among high-innovation forecasts, analyst characteristics still contribute significantly to explaining return responses, but return responses to analyst forecast revisions are inconsistent with rational expectations of forecast accuracy.

Table 4 shows that for low-innovation forecasts, the results for the forecast accuracy model are similar to the full sample results and to the results for high-innovation forecasts in Table 3. However, in the unconstrained stock returns model, only the forecast frequency coefficient is statistically significant (at the 10 percent level). More importantly, a test of the joint significance of all the analyst characteristics in the unconstrained returns model

---

[19] The simple correlation between prior uncertainty and forecast horizon is 0.66, indicating that longer forecast horizons are strongly associated with more prior uncertainty among analysts.

[20] We repeated the analysis including prior uncertainty in both the forecast accuracy model and the returns model. We find that prior uncertainty and forecast horizon are both associated with reduced forecast accuracy in the forecast accuracy model. In the returns model, prior uncertainty is strongly associated with increased return responses to forecast revisions, but forecast horizon is associated with reduced return responses to forecast revisions, consistent with the forecast accuracy model. Tests of the forecast and returns model coefficients show that these results are still inconsistent with investors' use of rational expectations of forecast accuracy in responding to analyst forecast revisions.

TABLE 3
Regression Model Coefficient Estimates for the Forecast Accuracy and Returns Models (high-innovation forecast revisions:[a] n = 25,445)

$$
\begin{aligned}
ACCURACY_{ijt} = {} & \alpha_0 + \alpha_1 DAYS\_ELAPSED_{ijt} + \alpha_2 FOR\_HORIZON_{ijt} \\
& + \alpha_3 LAG\_ACCURACY_{ijt} + \alpha_4 BROKER\_SIZE_{ijt} \\
& + \alpha_5 FOR\_FREQUENCY_{ijt} + \alpha_6 FIRM\_EXPERIENCE_{ijt} \\
& + \alpha_7 COMPANIES_{ijt} + \alpha_8 INDUSTRIES_{ijt} + \varepsilon_{ijt} \qquad (3)
\end{aligned}
$$

$$
\begin{aligned}
CAR_{ijt} = {} & \gamma_0 + \gamma_1 REVP_{ijt} \times (\alpha_0^* + \alpha_1^* DAYS\_ELAPSED_{ijt} + \alpha_2^* FOR\_HORIZON_{ijt} \\
& + \alpha_3^* LAG\_ACCURACY_{ijt} + \alpha_4^* BROKER\_SIZE_{ijt} \\
& + \alpha_5^* FOR\_FREQUENCY_{ijt} + \alpha_6^* FIRM\_EXPERIENCE_{ijt} \\
& + \alpha_7^* COMPANIES_{ijt} + \alpha_8^* INDUSTRIES_{ijt}) + \varepsilon_{ijt} \qquad (4)
\end{aligned}
$$

Panel A: Parameter Estimates

	Forecast Model Coefficients			*Returns Model Coefficients*		
	Parameter	*Estimate*	*Asymptotic Std. Error*	*Parameter*	*Estimate*	*Asymptotic Std. Error*
DAYS_ELAPSED	α_1	−0.036***	0.0061	α_1^*	0.287***	0.0567
FOR_HORIZON	α_2	−0.238***	0.0066	α_2^*	0.292***	0.0608
LAG_ACCURACY	α_3	0.067***	0.0062	α_3^*	0.087	0.0577
BROKER_SIZE	α_4	0.027***	0.0065	α_4^*	0.374***	0.0605
FOR_FREQUENCY	α_5	0.056***	0.0066	α_5^*	−0.015	0.0627
FIRM_EXPERIENCE	α_6	0.016***	0.0063	α_6^*	0.014	0.0578
COMPANIES	α_7	−0.032***	0.0079	α_7^*	0.012	0.0737
INDUSTRIES	α_8	−0.017**	0.0078	α_8^*	−0.019	0.0719
REVP[b]				γ_1	3.557***	0.1360

Panel B: Tests of Equality of Forecast and Returns Model Coefficients

	Parameters	*Likelihood Ratio Statistic*	*Marginal Significance Level*
DAYS_ELAPSED	$\alpha_1 = \alpha_1^*$	12.73	<0.001
FOR_HORIZON	$\alpha_2 = \alpha_2^*$	23.29	<0.001
LAG_ACCURACY	$\alpha_3 = \alpha_3^*$	0.67	0.412
BROKER_SIZE	$\alpha_4 = \alpha_4^*$	18.83	<0.001
FOR_FREQUENCY	$\alpha_5 = \alpha_5^*$	1.23	0.271
FIRM_EXPERIENCE	$\alpha_6 = \alpha_6^*$	0.65	0.424
COMPANIES	$\alpha_7 = \alpha_7^*$	3.92	0.048
INDUSTRIES	$\alpha_8 = \alpha_8^*$	0.57	0.447
Overall (joint) test	$\alpha_i = \alpha_i^*$, i = 1,...,8	75.43	<0.001

---

*, **, *** Significantly different from zero at the 10, 5, and 1 percent levels, respectively.

[a] High-innovation forecast revisions are above both the analyst's own prior forecast and the consensus forecast, or else are below both forecast revisions.

[b] We estimate γ_1 when we constrain the α_i^* coefficients in Equation (4) to equal the corresponding α_i coefficients from Equation (3).

Variables are defined in Tables 1 and 2.

On the other hand, we find no consistent evidence of either a lead or lag association between timely-reviewed earnings (or unexpected earnings) and returns. For timely-reviewed earnings, the association between earnings and returns appears to be largely contemporaneous. Collectively, our results imply that timely-reviewed quarterly earnings include significantly more of the economic information that is embedded in concurrent stock returns.

The next section reviews related research. Section III presents our models and discusses the hypotheses. Section IV describes empirical measures and sample selection criteria. Section V presents results and Section VI concludes the study.

II. BACKGROUND AND RELATED LITERATURE

SEC rules on quarterly filings require firms to disclose sales, gross profit, net income, and earnings per share (Regulation S-K, item 302(a), SEC [1999b]). The SEC does not require firms to have this information audited, but does require that it be reviewed.[2] Before the SEC required all firms to have their auditor review the firm's quarterly reports on a timely basis, it permitted firms to include a review letter with their filing, but only when the review occurred prior to the filing (a timely review). Ettredge et al. (1994), however, find that for their sample of SEC registrants, only 8 percent included the review letter with their filings.

Since our primary research objective is to determine whether reported quarterly earnings better reflect the economic information impounded in stock returns when the auditor reviews those earnings on a timely basis, our research design does not require the securities markets to know the timing of the interim review. Ettredge et al. (1994) argue, however, that securities markets become aware of review timing through informal channels. They support this claim by providing evidence that systematic differences exist between the characteristics of firms that choose timely reviews and those firms that choose retrospective reviews. Specifically, they show that firms purchasing timely reviews have higher asset levels, are more highly leveraged, have a smaller proportion of management ownership, have recently issued new securities, exhibit a higher proportion of receivables to total assets, and employ an internal audit staff to review quarterly information prior to public release. Raedy and Rock (2001) find that markets react more strongly to unexpected quarterly earnings when the review is timely. Together, Ettredge et al. (1994) and Raedy and Rock (2001) provide evidence that market participants act as if they are aware of the timing of reviews.

We expect timely reviews to increase the quality of reported interim earnings by reducing unintentional estimation errors, and by curbing earnings management. With respect to estimation errors, interim earnings require more estimation than annual earnings, because deferrals and accruals at the end of each quarter result from judgments concerning anticipated results for the rest of the fiscal year (AICPA 1972).[3] With respect to earnings management, Watts and Zimmerman (1986, Chapter 9) describe the incentives management has to manipulate income, and Mendenhall and Nichols (1988) note that management has more opportunity to manipulate interim earnings when the earnings reports are unaudited. If auditors review interim earnings on a timely basis, then the increased auditor involvement throughout the year would likely curb management's ability to manipulate interim earnings.

---

[2] Auditors generally limit reviews to inquiries and analytical procedures (SAS No. 71, AICPA 1972).

[3] Rangan and Sloan (1998) note that this feature of the integral approach (AICPA 1973) results in intra-year dependencies in quarterly earnings. We address this issue by aligning cross-sectionally the specific quarters in the fiscal year.

Empirical evidence is consistent with timely reviews reducing estimation errors and curbing earnings management. Givoly and Ronen (1981) report evidence of fourth-quarter "smoothing" adjustments, and Collins et al. (1984) and Bathke and Lorek (1984) describe fourth-quarter adjustments as "settling-up." The BRC (1999, 35) blames the incidence of fourth-quarter adjustments on "inaccuracies not detected during the preceding three quarters," and further claims that "increased involvement by the outside auditors and the audit committee in the interim financial reporting process should result in more accurate interim reporting." Consistent with the BRC's claim that timely-reviewed interim earnings are of higher quality, Ettredge et al. (2000) show that firms choosing timely reviews record fewer fourth-quarter adjustments than do firms choosing retrospective reviews.

In summary, prior and concurrent research indicates that timely reviews lead to fewer fourth quarter adjustments than do retrospective reviews, and that the return response to interim-quarter unexpected earnings is stronger when the review is timely rather than retrospective. Together, these findings suggest that timely reviews enhance the quality of interim earnings. Earnings quality, however, is unobservable, such that the evidence that timely reviews enhance earnings quality is necessarily indirect. Our study provides additional indirect evidence that timely reviews enhance interim-quarter earnings quality by documenting that the earnings-returns relation is more contemporaneous for firms that purchase timely reviews than for firms that do not.

III. MODEL AND HYPOTHESIS DEVELOPMENT

Our research objective is to determine whether timely-reviewed quarterly earnings better reflect the economic information contained in contemporaneous stock returns than do retrospectively-reviewed quarterly earnings. Stock returns measure the change in a firm's (perceived) value over a period of time, and thus returns serve as a measure of the firm's economic earnings. Because we are interested in whether timely- vs. retrospectively-reviewed accounting earnings differ in the extent to which they reflect the information impounded in contemporaneous stock returns (a measure of economic earnings), we relate returns to earnings realizations (rather than earnings innovations), similar to Dechow (1994).

We are interested in whether quarterly earnings reflect the same economic information that is embedded in returns during the earnings period. We therefore match the returns cumulation period to the fiscal quarter. Our model is thus:

$$RET_q = \alpha_{i,0} + \alpha_{i,1} NI_q + \varepsilon_q \tag{1}$$

where:

RET_q = raw stock returns cumulated during fiscal quarter q, $q \in \{1,2,3,4\}$;
NI_q = firm earnings for fiscal quarter q, standardized by firm market value of equity at the beginning of the period;
$\alpha_{i,k}$ = coefficients, $i \in \{T,R\}$ for either timely-reviewed or retrospectively-reviewed firms, and $k \in \{0,1\}$; and
ε_q = error term.

We initially estimate Equation (1) for each of the four fiscal quarters, using a quarterly returns window. If timely reviews lead to reported earnings that more accurately reflect the economic information already impounded in returns during the quarter, we will find a stronger association between quarterly returns and interim-quarter earnings for firms choosing timely reviews than for those firms choosing retrospective reviews, as stated in our first hypothesis (in alternative form) below:

> **H1:** (Interim-Quarter Hypothesis) The relation between quarterly returns and interim-quarter earnings is stronger in the subsample of timely-reviewed firms than in the subsample of retrospectively-reviewed firms ($\hat{\alpha}_{T,1} > \hat{\alpha}_{R,1}$).

Since firms report interim earnings as both independent quarterly performance measures and integral components of annual earnings (Rangan and Sloan 1998), we test our Interim-Quarter Hypothesis two ways. First, we use a one-quarter window and examine each of the four fiscal quarters independently. We then repeat our tests after expanding the window to include prior quarters of the fiscal year. That is, for the second quarter, we first test the relation between second-quarter earnings and second-quarter returns. We then repeat the tests after aggregating first- and second-quarter reported earnings (and returns), creating a two-quarter window. Similarly, for the third quarter, we first test the relation between third-quarter earnings and third-quarter returns, and then repeat the tests using a three-quarter window. Our expanding windows approach is similar to Easton et al. (1992), except that they aggregate multiple years, while we aggregate multiple quarters.

Hypothesis 1 posits that there is a stronger contemporaneous association between timely-reviewed interim earnings and returns than between retrospectively-reviewed interim earnings and returns. Respondents to SEC Release No. 34-42266 (SEC 1999a), however, argue that any short-run difference observed in interim earnings across review type would not persist with annual earnings, because all firms receive an annual audit. This suggests that there is a lagged association between retrospectively-reviewed interim earnings and returns. A lagged association would arise if either (1) firms report at least some of the economic information embedded in stock returns in later-quarters' earnings when auditors do not review interim earnings on a timely basis, or (2) the securities markets' reaction to information in interim earnings is not complete until the auditor reviews this information.

Although we develop and test these two explanations as independent hypotheses, they are not mutually exclusive. The first explanation posits that managers delay recognizing in earnings, certain transactions, or adjustments, until auditor scrutiny requires them to do so (even though market participants are aware of the transactions or adjustments). The second explanation posits that the market's reaction to new information is not complete until auditor involvement enhances the information's credibility.

The first explanation implies that with retrospective reviews, quarterly earnings issued later in a fiscal year contain previously omitted economic information that is already reflected in stock returns. The differences across review types occur because timely-reviewed interim earnings contain adjustments, such as estimates of bad debts or inventory corrections, that retrospectively-reviewed interim earnings do not (SEC 1999a; Ettredge et al. 2000). Firms may postpone recording correcting adjustments due to uncertainty about the interim accounting estimates or due to intentional earnings management (Mendenhall and Nichols 1988). In either case, however, quarterly earnings issued later in the same fiscal year either correct for the absence of prior adjustments, or the year-end audit requires these corrections.

To test whether firms report at least some of the economic information embedded in stock returns in later-quarters' earnings when auditors do not review interim earnings on a timely basis, we modify Equation (1) to include quarterly earnings issued later in the fiscal year, as follows:

$$RET_q = \beta_{i,0} + \beta_{i,1}NI_q + \beta_{i,2}\sum_{j=1}^{4-q} NI_{q+j} + \eta_q \tag{2}$$

where:

$\beta_{i,j}$ = coefficients, $i \in \{T,R\}$, $j \in \{0,1,2\}$; and
η_q = error term.

All other variables are as defined previously. The summation applies to reported earnings in the remaining quarters (q) of a given fiscal year. If stock returns lead retrospectively-reviewed earnings to a greater extent than they lead timely-reviewed earnings, then we should observe a stronger association between returns and future earnings in Equation (2) for retrospectively-reviewed firms than for timely-reviewed firms, as stated in our second hypothesis (in alternative form) below:

H2: (Returns-Lead-Earnings Hypothesis) The relation between quarterly returns and the quarterly earnings reported in subsequent quarters of the same fiscal year is stronger in the subsample of retrospectively-reviewed firms than in the subsample of timely-reviewed firms ($\hat{\beta}_{R,2} > \hat{\beta}_{T,2}$).

The second explanation for a stronger contemporaneous association between timely-reviewed interim earnings and returns than between retrospectively-reviewed interim earnings and returns is that stock prices react more quickly and completely to timely-reviewed earnings (even when both timely-reviewed and retrospectively-reviewed interim earnings contain new value-relevant information), perhaps because auditor involvement increases the credibility of timely-reviewed interim earnings. This explanation presumes that market participants can discern the timing of firms' reviews. As discussed previously, the evidence presented in Ettredge et al. (1994) and Raedy and Rock (2001) suggests that even if firms do not disclose review timing through formal or informal channels, market participants appear to infer the timing through underlying firm characteristics.

This explanation suggests that retrospectively-reviewed earnings lead stock returns. To test whether the market reacts more slowly to new information in retrospectively-reviewed interim earnings, we must model the return reaction to *unexpected* earnings information, which requires a fundamentally different approach from Equations (1) and (2) above. First, we must estimate unexpected earnings. Second, to identify the return reaction to the unexpected earnings information, the return cumulation window must encompass the release of that new earnings information.

Following Ohlson and Shroff (1992), we create a proxy for unexpected earnings by regressing current earnings on prior earnings. We begin by regressing quarterly earnings on earnings from the same quarter of the prior year, by review type:

$$NI_q = \delta_{i,0} + \delta_{i,1}NI_{q-4} + \mu_q \tag{3}$$

where:

$\delta_{i,j}$ = coefficients, $i \in \{T,R\}$, $j \in \{0,1\}$; and
μ_q = regression residual, which serves as our proxy for unexpected earnings.

All other variables are as defined previously. After estimating Equation (3) cross-sectionally, we regress quarterly returns on our unexpected earnings proxy (i.e., the regression residuals) from Equation (3), again by review type:

$$\mathrm{RET}_q = \gamma_{i,0} + \gamma_{i,1}\hat{\mu}_q + \gamma_{i,2}\sum_{j=1}^{q-1}\hat{\mu}_{q-j} + \nu_q \tag{4}$$

where:

$\gamma_{i,j}$ = coefficients, $i \in \{T,R\}$, $j \in \{0,1,2\}$; and
ν_q = error term.

All other variables are as defined previously. For this model, we cumulate returns from the day after the announcement of quarter $q - 1$ earnings through the quarter q earnings announcement date to ensure that RET_q fully reflects any information associated with unexpected earnings. We regress returns not only on the current-quarter's unexpected earnings ($\hat{\mu}_q$), but also on the sum of unexpected earnings from the preceding quarters for that same fiscal year. Thus, $\Sigma\ \hat{\mu}_{q-j}$ serves as our test metric for the second explanation, that interim earnings contain information about economic events that securities markets react to incompletely until the auditors review the information.

If this explanation applies, then retrospectively-reviewed earnings will lead stock returns to a greater extent than will timely-reviewed earnings. We would thus expect a larger coefficient on prior-period unexpected earnings in Equation (4) for retrospectively-reviewed firms than for timely-reviewed firms as stated in our third hypothesis (in alternative form) below:

H3: (Earnings-Lead-Returns Hypothesis) The relation between quarterly returns and the quarterly earnings reported in previous quarters of a fiscal year is stronger in the subsample of retrospectively-reviewed firms than in the subsample of timely-reviewed firms ($\hat{\gamma}_{R,2} > \hat{\gamma}_{T,2}$).

IV. SAMPLE SELECTION AND DESCRIPTIVE STATISTICS

Sample Selection Criteria and Industry Distribution

We contacted 1,025 firms listed in the 1995 *Directory of Corporate Affiliations*, and requested information about the timing of their interim reviews for the period 1990–1995. Of those 1,025 firms, 500 (48.8 percent) responded to our request. As shown in Table 1, we deleted 73 firms from regulated industries (utilities and banking), where governmental auditors, compliance officers, and others perform interim financial reviews beyond those performed by the independent auditor. We also deleted ten firms that refused to disclose the timing of their reviews, and five firms that changed from retrospective to timely reviews during the period. Of the remaining 412 firms disclosing review type, 328 (80 percent) reported that auditors reviewed their interim earnings on a timely basis. The remaining 84 firms (20 percent) reported that auditors reviewed their interim earnings at year-end (retrospective reviews).

Over our sample period, this yielded 2,472 firm-years composed of 1,968 firm-years with timely reviews and 504 firm-years with retrospective reviews. To be included in the final sample, each firm had to have sufficient data for our hypothesis tests, as well as for our sensitivity tests. We thus required eight consecutive quarters of data on average common

TABLE 1
Sample Selection and Screening Procedures—for the Full Sample and by Review Type

	Full Sample	*Timely Review Subsample*		*Retrospective Review Subsample*	
	Observations	*Observations*	*% of Sample*	*Observations*	*% of Sample*
Firms contacted through the *Directory of Corporate Affiliates, 1995*	1,025				
Less Firms:					
Not responding to inquiry	525				
From regulated industries	73				
Refusing to disclose timing	10				
Changing their review type	5				
Subtotal	412	328	80	84	20
Possible firm-years, 1990–1995	2,472	1,968		504	
Less firm-years:					
Missing data on Compustat	1,218	966		252	
Missing data on CRSP	130	112		18	
Subtotal	1,124	890	79	234	21
Less firm-years:					
With influential observations	3	3	—	—	—
Total firm-years in final sample	1,121	887	79	234	21

shares outstanding, earnings announcement date, and income before extraordinary items from Compustat. We also required annual Compustat data on accounts receivable, common stock, foreign assets, long-term debt, and total assets, and CRSP data on quarterly raw stock returns, announcement-day stock prices, and quarter-end stock prices.

As detailed in Table 1, these selection criteria resulted in 1,124 firm-years, 890 firm-years (79 percent) with timely reviews and 234 (21 percent) with retrospective reviews. We deleted three firm-years from our sample (all timely) that the DFFIT test identified as unduly influencing our results (see Neter et al. 1989). To further reduce the influence of extreme values, we also winsorize all accounting data at 1 percent (each tail). Our final sample consists of 1,121 firm-years, 887 (79 percent) timely and 234 (21 percent) retrospective.

Table 2 provides the sample distribution by industry. We report separate distributions for the full sample and the two subsamples. The largest industry group in our sample is the Steel and Machinery group, which comprises 28.01 percent of our total sample, 28.75 percent of the timely review subsample, and 25.22 percent of the retrospective review subsample. The comparability of these proportions makes it unlikely that the Steel and Machinery group affect differences between subsamples. In contrast, the Food and Wholesale industry groups collectively represent 11.39 percent of the timely review subsample, but comprise only 2.56 percent of the retrospective review subsample. Sensitivity analyses reported later indicate, however, that our inferences are not attributable to differences in the industry composition of the timely vs. retrospective review subsamples.

TABLE 2
Industry Distribution of Sample Firm-Years—for the Full Sample and by Review Type

		Full Sample		*Timely Review Subsample*			*Retrospective Review Subsample*		
Industry	*SIC Codes*	*Firm-Years*	*% of Sample*	*Firm-Years*	*% of Industry*	*% of Subsample*	*Firm-Years*	*% of Industry*	*% of Subsample*
Mining	1000–1499	66	5.89	54	81.82	6.09	12	18.18	5.13
Construction	1500–1999	24	2.14	6	25.00	0.68	18	75.00	7.69
Food	2000–2199	72	6.42	66	91.67	7.44	6	8.33	2.56
Textile	2200–2399	48	4.28	42	87.50	4.73	6	12.50	2.56
Wood	2400–2599	30	2.68	12	40.00	1.35	18	60.00	7.69
Paper and Printing	2600–2799	72	6.42	54	75.00	6.09	18	25.00	7.69
Chemicals	2800–2999	144	12.85	102	70.83	11.50	42	29.17	17.95
Plastic, Glass, and Cement	3000–3299	56	5.00	47	83.93	5.30	9	10.72	3.85
Steel and Machinery	3300–3999	314	28.01	255	81.21	28.75	59	18.79	25.22
Transportation	4000–4899	48	4.28	42	87.50	4.73	6	12.50	2.56
Wholesale	5000–5199	35	3.12	35	100.00	3.95	—	0.00	0.00
Retail	5200–5999	142	12.67	124	87.32	13.98	18	12.68	7.69
Other Services	7000–8999	70	6.24	48	68.57	5.41	22	31.43	9.41
Totals		1,121	100.00	887		100.00	234		100.00

Descriptive Statistics

Table 3 presents descriptive statistics for the full sample, along with the timely and retrospective subsamples. Our statistics show that when aggregated over a full year, neither stock returns nor earnings per share (0.168 and 0.175 for mean returns and $0.875 and $0.886 for mean earnings per share, for the timely review and retrospective review subsamples, respectively) significantly differ across the two subsamples. Additionally, mean stock prices do not significantly differ across the two subsamples (mean values of $23.41 and $21.49 for the timely and retrospective review subsamples, respectively). The earnings-to-price ratio, however, is significantly greater for the subsample of retrospectively-reviewed firms (mean values of 0.056 and 0.071 for the timely and retrospective subsamples, respectively). Thus, while neither the numerator nor denominator differs significantly across subsamples, together they create a significant difference in the E/P ratio across the subsamples.

Our statistics by quarter also show that quarterly earnings per share do not significantly differ across the two subsamples. However, mean quarterly earnings are lower for both review types in the first quarter and in the fourth quarter. Mean quarterly returns do not differ significantly across the two subsamples in three of the four quarters. Mean quarterly returns are lower and mean earnings per share are higher for retrospectively-reviewed firms in the second quarter than they are for either subsample in any other quarter.

Table 3 also reports descriptive statistics on variables that Ettredge et al. (1994) identify as varying systematically across the two review types. Consistent with Ettredge et al. (1994), we find that average total assets are larger, and the average percentage of common equity owned by management is smaller, for timely-reviewed firms.[4] The other characteristics appear comparable across subsamples.

V. RESULTS

Tests of the Interim-Quarter Hypothesis (H1)

Table 4 presents the results from tests of our first hypothesis. We predict that the contemporaneous relation between quarterly returns and interim-quarter earnings is stronger in the subsample of timely-reviewed firms than in the subsample of retrospectively-reviewed firms. For each quarter in Panel A, the quarterly returns window compounds returns over that fiscal quarter.

Panel A of Table 4 shows that the coefficients on timely-reviewed quarterly earnings are significantly positive in all quarters (coefficients of 1.26, 1.49, 0.96, and 0.86 for quarters 1 through 4, respectively; all p-values of 0.001). The coefficients on retrospectively-reviewed earnings are significantly weaker, or not significantly different from zero in the first two quarters, but are significantly positive in quarters 3 and 4 (coefficients of 0.50, 0.34, 0.93, and 1.00; p-values of 0.084, 0.206, 0.001, and 0.001 for quarters 1 through 4, respectively). The coefficients on timely-reviewed and retrospectively-reviewed earnings in quarters 3 and 4 are not significantly different from each other (p-values of 0.465 and 0.455).[5] However, the coefficients on retrospectively-reviewed earnings are significantly smaller than the coefficients on timely-reviewed quarterly earnings in quarters 1 and 2 (p-values of 0.020 and 0.008, respectively). Thus, our results support H1 for the first two interim quarters.

Table 4, Panel B presents our results from testing H1 using expanded windows that increase progressively over the quarters, such that the second-quarter window equals the

---

[4] Section V reports the results of sensitivity tests that control for differences in firm characteristics.

[5] We estimate the model for both review types simultaneously. Thus, this inference is based on a standard t-test of linear restrictions that $\hat{\alpha}_{T,1} = \hat{\alpha}_{R,1}$.

TABLE 3
Sample Descriptive Statistics and Tests for Differences across Timely Review and Retrospective Review Subsamples

	Full Sample (n = 1,121)			*Timely Review Subsample (n = 887)*			*Retrospective Review Subsample (n = 234)*			
Variable	*Mean*	*Median*	*Standard Deviation*	*Mean*	*Median*	*Standard Deviation*	*Mean*	*Median*	*Standard Deviation*	*Differences in Subsample Means*
Annual Returns	0.169	0.137	0.293	0.168	0.135	0.298	0.175	0.140	0.274	−0.007
Annual Earnings	0.877	0.715	1.326	0.875	0.755	1.259	0.886	0.595	1.575	−0.011
Stock Price	23.013	20.875	13.534	23.410	22.177	13.237	21.486	16.249	14.548	1.924
Earnings/Price Ratio	0.059	0.062	0.056	0.056	0.061	0.054	0.071	0.071	0.063	−0.015***
Quarterly Returns										
Qtr. 1	0.054	0.037	0.158	0.051	0.035	0.161	0.064	0.019	0.146	−0.013
Qtr. 2	0.026	0.017	0.144	0.030	0.023	0.146	0.012	0.003	0.133	0.018*
Qtr. 3	0.038	0.010	0.119	0.036	0.011	0.115	0.045	0.006	0.131	−0.009
Qtr. 4	0.056	0.024	0.130	0.056	0.022	0.130	0.058	0.028	0.129	−0.002
Quarterly Earnings										
Qtr. 1	0.198	0.138	0.354	0.202	0.148	0.325	0.180	0.119	0.456	0.022
Qtr. 2	0.264	0.190	0.381	0.260	0.198	0.347	0.283	0.160	0.494	−0.023
Qtr. 3	0.242	0.195	0.433	0.239	0.200	0.418	0.256	0.163	0.489	−0.016
Qtr. 4	0.194	0.201	0.547	0.189	0.210	0.547	0.216	0.174	0.550	−0.027

(continued on next page)

TABLE 3 (continued)

Variable	*Full Sample (n = 1,121)* *Mean*	*Median*	*Standard Deviation*	*Timely Review Subsample (n = 887)* *Mean*	*Median*	*Standard Deviation*	*Retrospective Review Subsample (n = 234)* *Mean*	*Median*	*Standard Deviation*	*Differences in Subsample Means*
Total Assets (in millions)	2,484	768	4,277	2,778	810	4,677	1,388	651	1,860	1,390***
Long-Term Debt to Total Asset Ratio	0.260	0.264	0.196	0.264	0.273	0.191	0.245	0.219	0.216	0.019
Foreign Assets to Total Asset Ratio	0.268	0.000	0.332	0.274	0.000	0.331	0.244	0.000	0.338	0.030
Receivables to Total Asset Ratio	0.165	0.151	0.115	0.166	0.151	0.117	0.162	0.148	0.109	0.004
New Security Issues	0.022	0.000	0.076	0.023	0.000	0.076	0.019	0.000	0.077	0.004
Mgt. Ownership Percentage	0.155	0.126	0.161	0.126	0.069	0.167	0.162	0.152	0.159	−0.036***

---

*, ** and *** Signify two-tailed significance at the 10, 5 and 1 percent levels, respectively.

We define the variables as follows:

Returns = raw stock returns for quarter q, cumulated over the period beginning the first day of fiscal quarter q, through the last day of fiscal quarter q. Annual returns cumulate returns over quarters 1 through 4;
Earnings = earnings per share reported for fiscal quarter q. Annual earnings cumulate earnings over quarters 1 through 4;
Stock Price = stock price per share at the end of the fiscal year;
Earnings/Price Ratio = annual earnings per share divided by stock price per share, measured at the beginning of the fiscal year;
Total Assets = total assets (in millions), measured at the beginning of the fiscal year;
Long-Term Debt to Total Asset Ratio = ratio of long-term debt to total assets, measured at the beginning of the fiscal year;
Foreign Assets to Total Asset Ratio = ratio of foreign assets to total assets, measured at the beginning of the fiscal year;
Receivables to Total Asset Ratio = accounts receivable as a percentage of total assets, measured at the beginning of the fiscal year;
New Security Issues = increase in common stock as a percentage of total assets, measured over the prior fiscal year; and
Mgt. Ownership Percentage = percentage of common equity owned by management, measured at the beginning of the fiscal year.

TABLE 4
Tests of H1: OLS Regressions of Returns on Current Earnings Levels—by Review Type

$$RET_q = \alpha_{i,0} + \alpha_{i,1}NI_q + \varepsilon_q$$

Panel A: Quarterly Returns Window Results

	Timely Review Subsample			*Retrospective Review Subsample*			*Test of Differences in Coefficients between Subsamples*	
Returns Interval	*Adjusted R^2*	*Intercept*	*NI_q*	*Adjusted R^2*	*Intercept*	*NI_q*	*Intercept*	*NI_q*
expected sign:		±	+		±	+	±	+
Qtr. 1	0.15	0.51 10.39 (0.001)	1.26 4.39 (0.001)	0.20	0.76 6.85 (0.001)	0.50 1.38 (0.084)	−0.25 −2.13 (0.033)	0.77 2.05 (0.020)
Qtr. 2	0.08	0.23 4.56 (0.001)	1.49 5.83 (0.001)	0.03	0.19 1.49 (0.137)	0.34 0.82 (0.206)	0.04 0.29 (0.774)	1.15 2.41 (0.008)
Qtr. 3	0.02	0.30 6.98 (0.001)	0.96 4.69 (0.001)	0.03	0.46 4.76 (0.001)	0.93 3.20 (0.001)	−0.16 −1.54 (0.124)	0.03 0.09 (0.465)
Qtr. 4	0.09	0.52 12.09 (0.001)	0.86 5.31 (0.001)	0.26	0.58 6.69 (0.001)	1.00 5.00 (0.001)	−0.06 −0.65 (0.518)	−0.14 −0.75 (0.445)

(continued on next page)

TABLE 4 (continued)

Panel B: Expanded Returns Window Results

	Timely Review Subsample			*Retrospective Review Subsample*			*Test of Differences in Coefficients between Subsamples*	
Returns Interval	*Adjusted R^2*	*Intercept*	*NI_q*	*Adjusted R^2*	*Intercept*	*NI_q*	*Intercept*	*NI_q*
expected sign:		±	+		±	+	±	+
Qtr. 1	0.15	0.51 10.39 (0.001)	1.26 4.39 (0.001)	0.20	0.76 6.85 (0.001)	0.50 1.38 (0.084)	−0.25 −2.13 (0.033)	0.77 2.05 (0.020)
Qtr. 1–2	0.25	0.89 16.21 (0.001)	1.62 11.16 (0.001)	0.24	0.85 5.35 (0.001)	1.02 3.34 (0.001)	0.04 0.26 (0.792)	0.60 1.76 (0.040)
Qtr. 1–3	0.26	0.73 8.78 (0.001)	1.51 9.79 (0.001)	0.21	0.62 2.73 (0.006)	0.95 3.16 (0.001)	0.11 0.46 (0.644)	0.56 1.65 (0.050)
Qtr. 1–4	0.34	0.99 7.00 (0.001)	1.54 9.77 (0.001)	0.34	1.00 3.55 (0.001)	1.15 3.88 (0.001)	−0.02 −0.05 (0.961)	0.39 1.17 (0.243)

Each cell in Panels A and B reports the OLS coefficient, the White's (1980) adjusted t-statistic and its associated p-value. We report one-tailed p-values for those tests associated with directional predictions and two-tailed p-values for tests associated with nondirectional predictions.

We define the variables as follows:

RET_q = raw stock returns for quarter q, cumulated over the period beginning the first day of fiscal quarter q, through the last day of fiscal quarter q. The expanded window through quarter 2 cumulates RET_q over quarters 1 and 2. The expanded window through quarter 3 cumulates RET_q over quarters 1 through 3. Annual returns cumulate RET_q over quarters 1 through 4; and

NI_q = earnings per share for quarter q. We deflate NI_q by the beginning of the period stock price. The expanded window through quarter 2 cumulates NI_q over quarters 1 and 2. The expanded window through quarter 3 cumulates NI_q over quarters 1 through 3. Annual earnings cumulate NI_q over quarters 1 through 4.

first- and second-quarter windows combined, and the third-quarter window equals the first-, second-, and third-quarter windows combined. The fourth-quarter window is thus an annual window, and the first-quarter window (presented for comparison) is the same as in Panel A.

As above, the coefficients on timely-reviewed earnings for quarters 2 through 4 are all significantly positive (coefficients of 1.62, 1.51, and 1.54 for quarters 2 through 4, respectively; all p-values of 0.001). The coefficients on retrospectively-reviewed earnings are also all significantly positive in quarters 2, 3 and 4 (coefficients of 1.02, 0.95, and 1.15; p-values of 0.001 for each quarter). Tests of differences between the coefficients in these expanded window regressions reveal that the coefficients on timely-reviewed earnings are significantly greater than the coefficients on retrospectively-reviewed earnings through the quarter 3 period (p-values of 0.040 and 0.050, for quarters 2 and 3, respectively). In contrast, we find no significant difference across the two subsamples in the coefficients on annual earnings (p-value of 0.243).

To summarize, the results in Panels A and B of Table 4 indicate that the associations between returns and contemporaneous retrospectively-reviewed interim-quarter earnings are insignificant, or significantly weaker than those for timely-reviewed interim-quarter earnings in the first two fiscal quarters. As the returns and earnings window increases, however, the significance of the association between returns and retrospectively-reviewed earnings increases. Over an annual window, the association between returns and retrospectively-reviewed earnings does not differ significantly from the association between returns and timely-reviewed earnings. As we discussed in Section III, if the contemporaneous association between interim earnings and returns is significantly stronger for timely-reviewed firms, but the associations between annual earnings and returns do not differ across review type, then the interim earnings-returns association for retrospectively-reviewed firms is largely noncontemporaneous.

This pattern of results suggests that for retrospectively-reviewed earnings, either returns lead quarterly earnings (H2), or quarterly earnings lead returns (H3), within a fiscal year. Either way, our evidence suggests that retrospectively-reviewed earnings capture economic information on an *annual* basis that is not always evident in the contemporaneous association between *quarterly* earnings and returns.

Tests of Returns Leading Earnings (H2)

The evidence in Table 4 supports H1; the contemporaneous relation between returns and earnings is stronger for timely-reviewed firms. Hypothesis 2 posits that the explanation for this result is that timely-reviewed earnings reflect, in a more timely manner, the information impounded in contemporaneous stock returns. In other words, for retrospectively-reviewed earnings, subsequent quarters correct previous-quarter earnings by capturing information already impounded in stock returns. In regressions of quarterly returns on current- and future-quarter earnings, H2 would therefore predict that the coefficient on future earnings is larger in the subsample of retrospectively-reviewed firms than in the subsample of timely-reviewed firms.

Table 5 shows that for timely-reviewed firms, the coefficients on future earnings are not significantly different from zero in the first two quarters, although the coefficient is significantly positive in quarter 3 (coefficients of 0.32, 0.39, and 0.44; p-values of 0.144, 0.110, and 0.021 for quarters 1 through 3, respectively).[6] For the subsample of retrospectively-reviewed firms, however, quarterly returns are significantly associated with future

---

[6] Because only quarters 1 through 3 need future quarter earnings to complete the fiscal year, we cannot test H2 for the fourth quarter.

TABLE 5

Tests of H2 (Returns Lead Earnings): OLS Regressions of Returns on Current and Future Earnings Levels—by Review Type

$$RET_q = \beta_{l,0} + \beta_{l,1} NI_q + \beta_{l,2} \sum_{j=1}^{4-q} NI_{q+j} + \eta_q$$

	Timely Review Subsample				*Retrospective Review Subsample*				*Test of Differences in Coefficients between Subsamples*		
Returns Interval	*Adjusted R^2*	*Intercept*	NI_q	$\sum_{j=1}^{4-q} NI_{q+j}$	*Adjusted R^2*	*Intercept*	NI_q	$\sum_{j=1}^{4-q} NI_{q+j}$	*Intercept*	NI_q	$\sum_{j=1}^{4-q} NI_{q+j}$
expected sign:		±	+	+		±	+	+	±	+	−
Qtr. 1	0.15	0.48	1.19	0.32	0.23	0.53	−0.38	1.75	−0.05	1.57	−1.43
		5.88	2.36	1.06		3.55	−0.62	2.69	−0.30	2.32	−2.07
		(0.001)	(0.009)	(0.144)		(0.001)	(0.733)	(0.004)	(0.762)	(0.010)	(0.019)
Qtr. 2	0.08	0.22	1.23	0.39	0.04	0.12	−0.04	1.03	0.10	1.27	−0.64
		3.06	3.17	1.22		0.73	0.10	1.97	0.57	2.43	−1.06
		(0.002)	(0.001)	(0.110)		(0.466)	(0.536)	(0.024)	(0.564)	(0.008)	(0.144)
Qtr. 3	0.03	0.07	0.77	0.44	0.06	−0.35	0.93	1.52	0.37	−0.86	−0.75
		1.19	2.33	2.03		−2.03	1.44	3.06	−2.31	−0.23	−2.07
		(0.233)	(0.010)	(0.021)		(0.043)	(0.075)	(0.001)	(0.021)	(0.590)	(0.019)

Each cell in Panels A and B reports the OLS coefficient, the White's (1980) adjusted t-statistic and its associated p-value. We report one-tailed p-values for those tests associated with directional predictions and two-tailed p-values for tests associated with nondirectional predictions.

We define the variable as follows:

RET_q = raw stock returns for quarter q, cumulated over the period beginning the first day of fiscal quarter q, through the last day of fiscal quarter q.

NI_q = earnings per share for quarter q. We deflate NI_q by the beginning of the period stock price.

$\sum NI_{q+j}$ = sum of earnings per share reported in subsequent quarters ($j \in (1, 4-q)$) within a fiscal year;

reported earnings in all three interim quarters (coefficients of 1.75, 1.03, and 1.52; p-values of 0.004, 0.024, and 0.001 for quarters 1 through 3, respectively). Our tests of coefficient differences show that the relations between quarterly returns and future reported earnings are significantly stronger in the subsample of retrospectively-reviewed firms than in the subsample of timely-reviewed firms, in the first and third quarters (p-values of 0.019 and 0.019, respectively).

Additionally, the current earnings results reported in Table 5 are consistent with those reported in Table 4. Specifically, timely-reviewed *current*-period earnings are significantly associated with returns, after controlling for future earnings, in all three interim quarters. In contrast, retrospectively-reviewed *current*-period earnings are significantly associated with returns, after controlling for future earnings, only in the third quarter (p-value of 0.075). As in Table 4, the coefficients on *current*-period earnings are significantly higher for timely-reviewed firms than for retrospectively-reviewed firms, in quarters 1 and 2.

The evidence presented in Table 5 suggests that returns lead earnings to a greater extent when the review is retrospective than when the review is timely. This evidence supports H2 and provides one explanation for why the differences in the associations between interim earnings and returns across review type do not persist through annual windows.

Tests of Earnings Leading Returns (H3)

Even if "returns lead earnings" (H2) for retrospectively-reviewed firms, because some of the information currently impounded in returns is not reflected in those earnings until subsequent quarters, it is also possible that "earnings lead returns" (H3) if the initial market reaction to (unexpected) retrospectively-reviewed earnings is incomplete. Thus, we also test H3: returns do not fully reflect information in retrospectively-reviewed quarterly earnings until later periods. This could occur if auditor involvement gives timely-reviewed earnings more credibility than retrospectively-reviewed earnings. The market reaction to unexpected earnings would therefore be comparatively faster for a timely-reviewed firm than for a retrospectively-reviewed firm (if the market is awaiting confirmation from the auditor). In regressions of quarterly returns on current- and previous-quarter unexpected earnings, H3 predicts that the coefficient on previous-quarter unexpected earnings is larger for retrospectively-reviewed firms than for timely-reviewed firms.

Panel A of Table 6 presents results from regressing current earnings on earnings from the same quarter of the prior year. The residuals from these regressions serve as our measure of quarterly unexpected earnings. We find significant associations between and NI_q and NI_{q-4} in all four quarters for the subsample of firms with timely reviews (estimated coefficients of 0.52, 0.53, 0.19, and 0.32 for quarters 1 through 4, respectively, all p-values of 0.001). We also find significant associations between NI_q and NI_{q-4} in the first, third, and fourth quarters for the subsample of firms with retrospective reviews, but not in the second quarter (estimated coefficients of 0.40, 0.04, 0.34, and 0.32, p-values of 0.001, 0.223, 0.001, and 0.001, for quarters 1 through 4, respectively).

Panel B of Table 6 presents results from our regressions of returns on current and past unexpected earnings. For our subsample of timely-reviewed firms, the coefficients on previous-quarter unexpected earnings are not significantly greater than zero in any of the three returns windows (p-values of 0.885, 0.260 and 0.998 for quarters 2 through 4, respectively).[7] We also find that the coefficients on previous-quarter unexpected earnings are not significantly greater than zero in any of the quarterly returns windows for our subsample

---

[7] We can test H3 only for quarters 2 through 4, because these tests examine the associations between quarterly returns and earnings reported in previous quarters of the fiscal year.

TABLE 6

Results from Two-Stage Estimation: (1) Estimating Unexpected Earnings, and (2) Tests of H3 (Earnings Lead Returns): Regressing Returns on the Unexpected Earnings Estimates for Current and Previous Periods—by Review Type

Panel A: Estimates of the Earnings Regression Residuals

$$NI_q = \delta_{l,0} + \delta_{l,1} NI_{q-4} + \mu_q$$

	Timely Review Subsample			*Retrospective Review Subsample*		
Returns Interval	*Adjusted R^2*	*Intercept*	*NI_{q-4}*	*Adjusted R^2*	*Intercept*	*NI_{q-4}*
expected sign:		±	+		±	+
Qtr. 1	0.33	0.01	0.52	0.14	0.01	0.40
		9.05	22.25		4.54	6.28
		(0.001)	(0.001)		(0.001)	(0.001)
Qtr. 2	0.29	0.01	0.53	0.00	0.02	0.04
		10.72	20.40		3.18	0.76
		(0.001)	(0.001)		(0.002)	(0.223)
Qtr. 3	0.04	0.02	0.19	0.20	0.02	0.34
		12.71	6.76		6.17	7.80
		(0.001)	(0.001)		(0.001)	(0.001)
Qtr. 4	0.10	0.01	0.32	0.08	0.02	0.32
		6.19	10.54		4.80	4.63
		(0.001)	(0.001)		(0.001)	(0.001)

(continued on next page)

TABLE 6 (continued)

Panel B: Tests of H3 (Earnings Lead Returns): OLS Regressions of Returns on Unexpected Current and Previous Earnings

$$RET_q = \gamma_{I,0} + \gamma_{I,1}\hat{\mu}_q + \gamma_{I,2}\sum_{j=1}^{q-1}\hat{\mu}_{q-j} + v_q$$

	Timely Review Subsample				*Retrospective Review Subsample*				*Test of Differences in Coefficients between Subsamples*		
Returns Interval	*Adjusted R^2*	*Intercept*	$\hat{\mu}_q$	$\sum_{j=1}^{q-1}\hat{\mu}_{q-j}$	*Adjusted R^2*	*Intercept*	$\hat{\mu}_q$	$\sum_{j=1}^{q-1}\hat{\mu}_{q-j}$	*Intercept*	$\hat{\mu}_q$	$\sum_{j=1}^{q-1}\hat{\mu}_{q-j}$
expected sign:		±	+	+		±	+	+	±	+	–
Qtr. 2	0.11	0.50	1.61	−0.35	0.06	0.45	0.14	−0.50	0.05	1.47	0.15
		10.03	5.31	−1.20		3.94	0.62	0.92	0.39	3.85	0.24
		(0.001)	(0.001)	(0.885)		(0.001)	(0.268)	(0.820)	(0.698)	(0.001)	(0.594)
Qtr. 3	0.03	0.20	0.88	0.11	−0.01	0.12	0.15	−0.01	0.08	0.73	0.12
		3.63	4.16	0.64		0.99	0.87	0.10	0.58	2.65	0.57
		(0.001)	(0.001)	(0.260)		(0.322)	(0.193)	(0.532)	(0.562)	(0.008)	(0.715)
Qtr. 4	0.20	0.76	0.32	−0.36	0.30	1.29	0.30	0.03	−0.53	0.02	−0.38
		12.72	2.43	−2.89		9.74	1.22	0.20	−3.63	0.06	−2.13
		(0.001)	(0.008)	(0.998)		(0.001)	(0.112)	(0.420)	(0.001)	(0.956)	(0.017)

Each cell in Panels A and B reports the OLS coefficient, the White's (1980) adjusted t-statistic and its associated p-value. We report one-tailed p-values for those tests associated with directional predictions and two-tailed p-values for tests associated with nondirectional predictions.

We define the variables as follows:

RET_q = raw stock returns for quarter q, cumulated over the period beginning the day after the announcement of quarter q−1 earnings through the quarter q earnings announcement date;

NI_q = earnings per share for quarter q. We deflate NI_q by the beginning of the period stock price;

$\hat{\mu}_q$ = residual from regressing quarterly earnings on earnings from the same quarter of the prior fiscal year, by review type, which serves as our proxy for unexpected earnings; and

$\sum \hat{\mu}_{q-j}$ = sum of earnings regression residuals ($j \in (1, q-1)$) within a fiscal year.

of retrospectively-reviewed firms (p-values of 0.820, 0.532, and 0.420 for quarters 2 through 4, respectively). Because H3 is only supported if the coefficient on prior-quarters' earnings is positive for retrospectively-reviewed firms, our evidence fails to support H3.[8] However, Panel B of Table 6 also shows that the contemporaneous association between returns and *current* unexpected earnings is significant for timely-reviewed firms in every quarter, but insignificant in every quarter for retrospectively-reviewed firms.

In summary, our results in Table 4 suggest that the association between returns and interim-quarter earnings is more contemporaneous when interim reviews are timely rather than retrospective. Quarterly earnings are significantly associated with quarterly returns for timely-reviewed firms in every fiscal quarter. Quarterly earnings are not strongly associated with returns for retrospectively-reviewed firms until fiscal quarter three (Table 4, Panel A). For retrospectively-reviewed firms, returns appear to lead earnings, which suggests that subsequent-quarter earnings correct previous-quarter earnings by capturing previously omitted information already impounded in stock returns (Table 5). Moreover, the contemporaneous association between returns and unexpected earnings is strong for timely-reviewed firms, but is insignificant for retrospectively-reviewed firms (Table 6). Overall, our results suggest that timely reviews lead to earnings that reflect economic information that is impounded in contemporaneous returns, while retrospective reviews fail to provide this discipline.

Additional Sensitivity Tests

Year-by-Year Regressions

The support we find for our first two hypotheses (Tables 4 and 5) is based on pooling the data across our six-year sample period (1990–1995).[9] Estimating Equations (1) and (2) by year and averaging the six year-by-year coefficients yields results substantially the same as those presented in the tables, again supporting our first two hypotheses. We pool the six year-by-year standard deviations in order to calculate the t-statistics that determine the level of significance.

When we retest H1 by year, we find that the average earnings coefficients for timely-reviewed earnings are significantly positive (p-values < 0.05) for each of the four quarters. The earnings coefficients for retrospectively-reviewed earnings are not significant for quarters 1 and 3, whereas the coefficients for quarters 3 and 4 are significant (p-values < 0.10, 0.05, respectively). As before, the average earnings coefficients are significantly larger for timely-reviewed interim earnings than for retrospectively-reviewed interim earnings in the first two quarters (p-values < 0.01).

When we expand the returns and earnings windows to two, three, and four quarters, our results are similar to those reported in Table 4, Panel B. The coefficient averages for timely-reviewed interim earnings cumulated through the second and third quarters are significantly larger than the coefficient averages for retrospectively-reviewed interim earnings (p-values < 0.05). As before, however, when we regress annual returns on annual earnings, the coefficients on earnings do not differ significantly between the timely review and retrospective review subsamples.

---

[8] To verify that the lack of support we find for H3 is not due to our specification of unexpected earnings, we retest our hypothesis, using the sum of the coefficients on earnings levels and changes as our proxy for unexpected earnings (Easton and Harris 1991; Ali and Zarowin 1992). This alternative model specification does not alter our conclusions; the coefficients on prior-quarter unexpected earnings are not significantly positive at conventional levels in any quarter for either subsample. Further, Freeman and Tse (1992) suggest that the relation between returns and unexpected accounting earnings is nonlinear. Including a nonlinear measure of unexpected earnings ($\hat{\mu}_q \times |\hat{\mu}_q|$) in our estimation of Equation (4), however, does not alter our inferences, and the coefficients on the nonlinear term are not significant.

[9] Sensitivity tests of H3 again fail to provide support for this hypothesis. We therefore limit the discussion of our additional tests to robustness checks of H1 and H2.

When we retest H2 by year, we find that the averages of the coefficients on future earnings for timely-reviewed firms are significant in the first and third quarters (p-values < 0.05) but not in the second quarter. The averages of the coefficients on future earnings for retrospectively-reviewed firms are significant (p-values < 0.01) for quarters 1 through 3, and, more importantly, significantly exceed the coefficient averages on future earnings for timely-reviewed firms (p-values < 0.01, 0.05, and 0.01, for quarters 1, 2 and 3, respectively). The year-by-year results suggest that returns lead earnings, in part, for both timely- and retrospectively-reviewed firms, but to a greater extent for retrospectively-reviewed firms.

Regressions Partitioned by Industries

Table 2 indicates that both the timely and retrospective review subsamples have a high concentration of firms in the Steel and Machinery industry. Additionally, two industry groups, Food and Wholesale, comprise over 11 percent of the timely review subsample. To assess the influence of possible industry effects on our tests of H1 and H2, we modify Equations (1) and (2), respectively, to include a dummy (0,1) indicator variable (IND) that represents one of these three industries, as follows:

$$\mathrm{RET}_q = \alpha_{i,0} + \alpha_{i,1}\mathrm{NI}_q + \alpha_{i,2}\mathrm{IND} + \alpha_{i,3}\mathrm{NI}_q\mathrm{IND} + \varepsilon_q \tag{5}$$

$$\begin{aligned}\mathrm{RET}_q = {} & \beta_{i,0} + \beta_{i,1}\mathrm{NI}_q + \beta_{i,2}\sum_{j=1}^{4-q}\mathrm{NI}_{q+j} + \beta_{i,3}\mathrm{IND} \\ & + \beta_{i,4}\mathrm{NI}_q\mathrm{IND} + \beta_{i,5}\sum_{j=1}^{4-q}\mathrm{NI}_{q+j}\mathrm{IND} + \eta_q.\end{aligned} \tag{6}$$

All other variables are as defined previously. The sum of the coefficients $\alpha_{i,1}$ and $\alpha_{i,3}$ in Equation (5) serves as our test metric for H1 and the sum of $\beta_{i,2}$ and $\beta_{i,5}$ in Equation (6) serves as our test metric for H2.

When we set IND to 1 for firms in the Steel and Machinery industry, we find that, with one exception, the results for firms in that industry are comparable to the results for the remaining firms in all other industries. In tests of H2, the coefficients on future earnings for timely-reviewed firms ($\beta_{T,2} + \beta_{T,5}$) in the Steel and Machinery industry are not significant in any of the interim quarters, similar to the results reported in Table 5. In contrast to Table 5, however, we find that the coefficients on future earnings for timely-reviewed firms ($\beta_{T,2}$) in industries other than Steel and Machinery are significantly positive for each of the three interim quarters (p-values < 0.10 for quarters 1 and 2 and p-value < 0.05 for quarter 3). Nonetheless, we still find that differences between the timely- vs. retrospectively-reviewed firms are significant in the first and third quarters for firms in the Steel and Machinery industry group, and in the first and second quarters for firms in other industries.

Similarly, when we set IND to 1 for firms in the Food or Wholesale industry groups, our inferences continue to hold, for both those firms in the Food and Wholesale industry groups, and for those firms in other industries. We therefore conclude that our inferences apply broadly, and are not specific to a few industries.

Regressions Partitioned by Firm Characteristics

Consistent with Ettredge et al. (1994), our Table 3 shows that average total assets are larger, and the average percentage of common equity owned by management is smaller, for timely-reviewed firms. Additionally, we find that the E/P ratio is larger for timely-reviewed firms. Of the total sample distribution based on total assets, we find that 51 percent of the timely-reviewed firms and 41 percent of the retrospectively-reviewed firms fall in the upper half of the distribution. Of the total sample distribution based on management ownership, 46 percent of the timely-reviewed firms and 69 percent of the retrospectively-reviewed firms fall into the upper half of the distribution. With respect to the E/P ratio, 47 percent of the timely-reviewed firms and 60 percent of the retrospectively-reviewed firms fall into the upper half of the distribution.

To test whether our results are consistent across large and small firms, as well as across firms with higher or lower management ownership and E/P ratios, we estimate Equations (5) and (6) by setting the indicator variable IND to 1 when a firm is in the upper half of the distribution for size, management ownership, and E/P, respectively.

With respect to firm size, we find that in both Equations (5) and (6), the coefficients on current earnings for timely-reviewed firms are again significantly positive (p-values < 0.01) for each of the four fiscal quarters, across both large and small firms. For retrospectively-reviewed firms, the coefficients on current earnings in Equations (5) and (6) are significantly positive (p-value < 0.05) only in the first quarter, and only for small firms. Similar to the main results reported in Table 4, the differences across review type are significant in quarters 1 and 2 (p-values < 0.10 and 0.01 for small firms, and p-values < 0.01 and 0.05 for large firms, for quarters 1 and 2, respectively) but are not significant in quarters 3 or 4. Retesting H2 also yields results similar to those reported in Table 5.

With respect to management ownership, we find no discernible differences between firms with high vs. low management ownership, in tests of H1 or H2. Turning to firms with high vs. low E/P ratios, tests of H1 yield substantially identical results to those reported in Table 4, except we also find that, for high E/P firms, the differences across review type are significant in all three interim quarters (p-values < 0.10, 0.05, and 0.10 for quarters 1, 2 and 3, respectively). For low E/P firms, the differences across review type are significant only in quarters 1 and 2 (p-values < 0.01 and 0.01 for quarters 1 and 2, respectively), similar to the results reported in Table 4. In tests of H2, the only difference we find for low E/P firms relative to the main results reported in Table 5 is that the coefficient on future earnings is significant for timely-reviewed firms in quarter 2 (p-value < 0.05). Collectively, the results from partitioning by firm size, management ownership, and E/P ratios provide evidence that our primary findings are not attributable to these firm-specific characteristics.

VI. CONCLUSION

We find a strong contemporaneous association between quarterly returns and interim-quarter earnings for firms that purchase timely reviews from their auditor. For firms with retrospective reviews, we find a significantly weaker association, or no association between quarterly returns and interim-quarter earnings in the first two quarters of the fiscal year. In contrast to these differences for interim-quarter earnings, we find no evidence that the associations between annual earnings and annual returns differ when the interim review is timely rather than retrospective. We investigate two possible explanations for this result: (1) retrospectively-reviewed firms' returns lead earnings, presumably because subsequent-quarter retrospectively-reviewed earnings correct previous-quarter earnings by capturing information already impounded in the firm's returns, and (2) retrospectively-reviewed firms' earnings lead the firm's returns, presumably because returns do not fully reflect information in retrospectively-reviewed quarterly earnings until later periods when the earnings have withstood auditor scrutiny.

Our evidence supports the first explanation—returns lead earnings to a greater extent for retrospectively-reviewed firms. We do not find that earnings lead returns for retrospectively-reviewed firms in any quarter. Moreover, the contemporaneous association between unexpected earnings and returns is significant for timely-reviewed firms in each quarter, but is not significant for retrospectively-reviewed firms. Collectively, our results suggest that timely reviews help ensure that a firm's quarterly earnings reflect economic information embedded in concurrent stock returns. These findings suggest that the recent SEC mandate requiring all firms to purchase timely reviews will likely increase the contemporaneous correlation between earnings and returns for firms that had been receiving retrospective reviews.

One limitation of our study arises from the possibility of sample selection bias. Our sample includes only those firms that responded to our inquiry regarding the timing of their interim reviews. If these respondents are not representative of the population, then our results would be biased. However, the comparability of our sample firms' characteristics to those reported by Ettredge et al. (1994) suggests that our sample is representative.

REFERENCES

Ali, A., and P. Zarowin. 1992. The role of earnings levels in annual earnings-returns studies. *Journal of Accounting Research* 30 (2): 286–296.

American Institute of Certified Public Accountants (AICPA). 1972. *Interim Financial Information.* Statement on Auditing Standards No. 71. New York, NY: AICPA.

———. 1973. *Interim Financial Reporting.* Accounting Principles Board Opinion No. 28. New York, NY: AICPA.

Bathke, A., Jr., and K. Lorek. 1984. The relationships between time-series models and the security market's expectation of quarterly earnings. *The Accounting Review* 59 (2): 163–176.

Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees (BRC). 1999. *Report and Recommendations of the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees.* New York, NY: New York Stock Exchange.

Collins, W., W. Hopwood, and J. McKeown. 1984. The predictability of interim earnings over alternative quarters. *Journal of Accounting Research* 22 (2): 467–479.

Dechow, P. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accruals. *Journal of Accounting and Economics* 18 (1): 3–42.

Easton, P., and T. Harris. 1991. Earnings as an explanatory variable for returns. *Journal of Accounting Research* 29 (1): 19–36.

———, ———, and J. Ohlson. 1992. Aggregate accounting earnings can explain most of securities returns: The case of long intervals. *Journal of Accounting and Economics* 15 (2/3): 119–142.

Ettredge, M., D. Simon, D. Smith, and M. Stone. 1994. Why do companies purchase timely quarterly reviews? *Journal of Accounting and Economics* 18 (2): 131–155.

———, ———, ———, and ———. 2000. The effect of the external accountant's review on the timing of adjustments to quarterly earnings. *Journal of Accounting Research* 38 (1): 195–207.

Freeman, R., and S. Tse. 1992. A nonlinear model of security price responses to unexpected earnings. *Journal of Accounting Research* 30 (2): 185–209.

Financial Executives Institute (FEI). 1999. *FEI Comments on the Recommendations of the Blue Ribbon Committee.* May 20. Morristown, NJ: FEI.

Givoly, D., and J. Ronen. 1981. "Smoothing" manifestations in fourth-quarter results of operations: Some empirical evidence. *Abacus* 17 (2): 174–193.

Mendenhall, R., and W. Nichols. 1988. Bad news and differential market reactions to announcements of earlier-quarters versus fourth-quarter earnings. *Journal of Accounting Research* 26 (Supplement): 63–86.

National Commission on Fraudulent Financial Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting.* Washington, D.C.: Government Printing Office.

Neter, J., W. Wasserman, and M. Kutner. 1989. *Applied Linear Regression Models*. 2nd edition. Homewood, IL: Richard D. Irwin.

Ohlson J., and P. Shroff. 1992. Changes versus levels in earnings as explanatory variables for returns: Some theoretical considerations. *Journal of Accounting Research* 30 (2): 210–226.

Raedy, K., and S. Rock. 2001. An empirical investigation of the impact of timely interim reviews: Evidence of earnings quality and market response behavior. Working paper, University of Colorado at Boulder.

Rangan, S., and R. Sloan. 1998. Implications of the integral approach to quarterly reporting for the post-earnings-announcement drift. *The Accounting Review* 73 (3): 353–371.

Securities and Exchange Commission (SEC). 1999a. *Final Rule: Audit Committee Disclosure, 17 CFR, Parts 210, 228, 229, and 240*. SEC Release No. 34–42266. Washington, D.C.: Government Printing Office.

———. 1999b. *Regulation S–K, item 302(a)*. 64 FR 73389, 73402, Dec. 30. Washington, D.C.: Government Printing Office.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (4): 817–838.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 275–296

The Effect of Information Complexity on Analysts' Use of That Information

Marlene A. Plumlee
University of Utah

ABSTRACT: In this study I investigate the relation between information complexity and financial analysts' use of that information. I rank by complexity six tax-law changes enacted by the Tax Reform Act of 1986, and then examine analysts' explicit forecasts of effective tax rates around those changes. I show that analysts' *revisions* of their forecasts of effective tax rates appear to impound the effects of the less complex tax-law changes but not the more complex changes. Furthermore, as expected, if analysts assimilate less complex (but not more complex) information, the magnitude of the *errors* in their forecasts of effective tax rates increases with the effects of the more complex tax-law changes, but is unrelated to the less complex changes. Taken together, these results indicate that analysts assimilate less complex information to a greater extent than they assimilate more complex information. Either analysts' abilities to incorporate specific information into their forecasts is a decreasing function of the complexity of that information, or analysts choose not to assimilate complex information because the cost would exceed the benefit. In either case, complexity reduces analysts' use of information. These results demonstrate the importance of considering information attributes, such as complexity, when investigating why analysts' forecasts fail to incorporate all public information.

Keywords: *information complexity; analysts' forecasts; information use; effective tax rate forecasts.*

Data Availability: *Data are available from public sources.*

I would like to thank the members of my dissertation committee (Douglas J. Skinner [Chair], Jeffery Aba banell, Phil Howrey, and Joel Slemrod), Darrell Brown, Martha Eining, Donna Philbrick, David Plumlee, Tayl Randall, Claire Richards, and especially Christine Botosan for their many insightful comments and suggestions. have also benefited from comments and suggestions from Ken Klassen, Jerry Salamon, three anonymous reviewe and the workshop participants at the University of Kansas, University of Massachusetts, University of Michiga University of Utah, Washington University, and the 1998 American Accounting Association Annual Meeting. I a grateful for the financial support of the Andersen Foundation and the Paton Research Fund of the University Michigan. This paper is based on my dissertation at the University of Michigan.

Submitted February 20
Accepted July 20

I. INTRODUCTION

Prior research finds that analysts fail to incorporate all available information into their earnings forecasts (e.g., Biddle and Ricks 1988; Bernard and Thomas 1990; Kim and Schroeder 1990; Abarbanell and Bernard 1992; Ali et al. 1992; Abarbanell and Bushee 1997, 1998). McEwen and Hunton (1999) and Chen et al. (2002) conjecture that information complexity reduces analysts' assimilation of the information, but these studies do not directly measure the relative complexity of the information to support this conjecture. In this study I test whether the complexity of information reduces analysts' use of that information. Specifically, I examine the extent to which analysts' effective tax rate (ETR) forecasts incorporate information related to six tax-law changes of varying complexity.

I measure the relative complexity of these six tax-law changes (enacted by the Tax Reform Act of 1986) based on tax professionals' responses to the American Institute of Certified Public Accountants' tax-complexity index questionnaire (AICPA 1993). I find that the analysts' *revisions* of their forecasts of ETRs appear to impound the effects of less complex tax-law changes but not the effects of more complex tax-law changes. In contrast, the magnitude of analysts' ETR forecast *errors* increases with the effects of more complex tax-law changes, but is unrelated to the effects of the less complex tax-law changes, as if analysts fully impound the less complex information into their ETR forecasts. Taken together, these results suggest that analysts impound less complex information more fully than they do more complex information, consistent with information complexity reducing their use of the tax-law change information.

Empirical evidence that the complexity of information reduces analysts' use of that information is relevant to standard setters and researchers. For example, the Financial Accounting Standards Board Mission Statement charges the Board to "promulgate standards only when the expected benefits exceed perceived costs" (FASB 2002, 2). The evidence presented in this study suggests that more complex information imposes a cost on even expert users, which in turn suggests that as the complexity of a standard increases, the expected benefit of that standard must also increase to justify promulgation.[1] In addition, the Securities and Exchange Commission (SEC) adopted the "plain English rule," which requires issuers to:

> use plain English principles in writing the front and back cover pages, summary and risk factor sections of prospectuses; revise current requirements for highly technical information in the front of prospectuses; and revise the rule on the preparation of prospectuses to provide companies with more specific guidance on the clarity required in the entire document. (SEC 1997, 1)

The goal of the "plain English rule" is to increase investors' understanding and help them make informed investment decisions. My evidence that the complexity of information reduces even expert analysts' use of information supports the SEC's adoption of the plain English rule because less sophisticated users are, by definition, less able than expert analysts to incorporate complex information into their investment decisions. Finally, researchers investigating the relevance of disclosed information should consider the complexity of that information when interpreting their results. For example, evidence that there is no significant relation between specific information and stock prices or other measures of information use does not necessarily mean that the information is irrelevant; it could also mean that the information is too complex to be used cost-effectively.

---

[1] In my study, complexity relates to the underlying economic changes resulting from a tax-law change, not to the application of an accounting standard. However, in both cases the complexity of the information imposes a cost on the users of that information

In the context of tax laws examined in this study, information complexity is a function of both rule complexity (understanding and applying a tax law to a static set of facts) and strategic complexity (understanding firms' tax-planning responses to new tax laws). Because my empirical measure of complexity ranks tax-law changes in terms of their rule complexity only, the validity of my inferences rests on whether this ranking reflects the overall information complexity of these tax laws. Section III provides evidence supporting the validity of my complexity ranking.

I organize the remainder of this paper as follows. Section II provides background information and develops the hypotheses in the context of prior research. In Section III, I discuss research design issues, including the models used, the complexity-ranking measure, and the sample. Section IV discusses the results of the empirical tests; Section V concludes the paper.

II. BACKGROUND AND HYPOTHESES

The Tax Reform Act of 1986 and Effective Tax Rates

The Tax Reform Act of 1986, formally adopted in October 1986, materially changed a number of specific tax-law provisions that affected corporate ETRs. This exogenous change provides a natural setting for examining the effect of complexity on analysts' forecasts. Some of the tax-law changes took effect immediately, while others became effective January 1, 1987. Accordingly, my analysis uses the forecasts of 1987 ETRs that analysts issued in the fourth quarter of 1986.

A firm's ETR equals its total income tax expense divided by its pretax accounting income. Total income tax expense is the product of taxable income and the tax rate applied to that income, reduced by any tax credits available to the firm. Tax-law changes typically affect ETRs through changes in total income tax expense stemming from changes in (1) the tax rates applied to taxable income; (2) the calculation of taxable income; or (3) the calculation or availability of tax credits. This study focuses on six tax-law changes included in the Tax Reform Act of 1986: two changes in tax rates (a decrease in the statutory tax rate and an increase in the capital gains tax rate); one change in tax rate and the calculation of taxable income (implementation of an alternative minimum tax); and three changes to tax credits (elimination of the investment tax credit; computation of the available research and development [R&D] tax credit; extensive changes to the foreign tax credit).

Tax-law changes do not directly affect pretax accounting income, the denominator of ETR. However, prior research documents firms' strategic responses to anticipated tax-law changes that could affect their ETRs through changes in pretax accounting income that do not proportionally affect total income tax expense (e.g., Gramlich 1991; Boynton et al. 1992; Collins and Shackelford 1992; Dhaliwal and Wang 1992; Manzon 1992; Scholes et al. 1992; Guenther 1994; Maydew 1997). I consider the potential implications of firms' strategic responses in my model (in Section III) and on the results (in Section IV).

Using ETR Forecasts

I use Value Line analysts' explicit forecasts of ETRs to examine how the complexity of the Tax Reform Act of 1986's tax-law changes affected analysts' use of that information. Analysts can more directly estimate the effect of tax-law changes on firms' ETRs than on firms' overall EPS, because each firm itemizes material items that directly affect its ETR in the annual report income tax footnote. I use these footnote data to estimate the effect of each tax-law change on the firm's ETR. Thus, testing the effect of the tax-law change on analysts' forecasts of the specific component of earnings most directly affected by the tax-law change (i.e., ETRs) is sharper than testing the effect of the tax-law change on overall earnings forecasts, which may be both directly and indirectly affected by the tax-law change. Provided that analysts exercise due care in forming ETR forecasts, using ETR forecasts increases the power of my tests and provides more reliable evidence on information used by analysts than I could obtain by analyzing their EPS forecasts.

I took several steps to assess whether Value Line analysts take care in forming their ETR forecasts. First, I carefully read the firm-specific text included with Value Line's numerical forecasts for evidence that analysts use tax-related information. During the third and fourth quarters of 1986, Value Line analysts frequently referred to the expected effect of proposed changes in the tax code on firm ETRs and earnings, including: (1) the direct effect of specific tax-law changes on ETRs; (2) the net effect of tax-law changes on earnings; (3) the offsetting effects of the tax-law changes on ETR and earnings; and (4) the effect of the tax-law changes on consumer and firm behavior.[2] These findings suggest that analysts carefully considered how the Tax Reform Act of 1986 tax-law changes would affect firms' ETRs and earnings. Second, I conducted empirical analyses that further support my assumption that Value Line analysts exercise care in developing their ETR forecasts. Specifically, using data described in Section II, I find that when an analyst forecasted an increase in a firm's ETR, she tended to forecast a decrease in the firm's related earnings; there is a negative correlation between the ETR forecasts revisions and earnings forecast revisions (untabulated correlation = −0.15; at p-value of 0.01). Moreover, when an analyst overestimated a firm's ETR, she tended to underestimate the firm's earnings; there is a negative correlation between the ETR forecast errors and earnings forecast errors (untabulated correlation = −0.20; at p-value of 0.001). These data suggest that analysts do take care in forming their ETR forecasts. Third, in personal interviews, two Value Line analysts assured me that they carefully consider the effects of intermediate forecasts (including ETRs) when forming future earnings forecasts.

Prior Literature and Hypothesis Development

Prior research documents that analysts apparently fail to impound all available information into their forecasts, but provides little insight into why (e.g., Biddle and Ricks 1988; Bernard and Thomas 1989; Kim and Schroeder 1990; Abarbanell 1991; Abarbanell and Bernard 1992; Ali et al. 1992; Abarbanell and Bushee 1997, 1998). More recent studies (Mikhail et al. 1997, 1999; Clement 1999) explore how *analyst* attributes or incentives contribute to variation in analysts' accuracy (which may stem from variation in their use of information, among other factors).[3] In contrast, I focus on information attributes, not analyst attributes.

Two streams of research suggest that an *information* attribute, complexity, may affect how efficiently market participants use that information. Theoretical and empirical judgment/decision-making research concludes that increased complexity of a task adversely affects judgment quality (e.g., Payne 1976; Einhorn et al. 1977; Iselin 1988; Paquette and

---

[2] For example, the Value Line report for Big B, Inc., on July 25, 1986 (prior to enactment of the Tax Reform Act of 1986), stated: **"Big B will benefit if the tax proposals now under consideration are enacted.** In fiscal 1987, the company's tax rate would decline considerably below our current estimates" (Value Line 1986a, 788) (bold in the original). The report for Prime Computer issued November 7, 1986 (after the enactment of the Tax Reform Act of 1986) stated: **"Our estimates include the effects of recent tax reform,** but we do not expect Prime's tax rate to be much affected by the new code. Though Prime will lose investment tax credits, research and development credits will likely produce a negligible net effect in this area" (Value Line 1986c, 1110) (bold in the original).

[3] Mikhail et al. (1997, 1999) conclude that forecast accuracy is important to analysts, and they document that analysts' forecasts become more accurate as analysts gain firm-specific experience. Clement (1999) finds that more experienced analysts and those who work for larger brokers tend to be more accurate, while analysts who follow larger numbers of firms and industries are less accurate.

Kida 1988; Payne et al. 1988).[4] This literature suggests that task complexity impairs judgment through decision-makers' strategy selection, where a strategy is the method or set of procedures an individual uses to incorporate information into decision making (e.g., expected utility maximization, satisficing, elimination by aspects). For example, Payne (1976) finds that, at a high level of task complexity, individuals use strategies that are analytically simpler to complete the task. Subsequent studies report similar findings (e.g., Payne 1982; Smith et al. 1982; Earley 1985; Bettman et al. 1990). These studies suggest that higher task complexity leads decision makers to adopt analytically simpler strategies that may result in incomplete use of available information.

A second stream of research suggests that information complexity likely impairs analysts' abilities to assimilate the information. For example, Hirst and Hopkins (1998) document that analysts fail to access comprehensive income information under certain reporting formats and suggest that "clear reporting" of information increases analysts' use of it. However, Hirst et al. (2002) hypothesize and find that analysts experienced with relevant comprehensive income information are unaffected by differences in reporting format.[5] In another experimental study, McEwen and Hunton (1999) document that analysts who forecast more accurately tend to emphasize different information than other analysts who forecast less accurately. McEwen and Hunton (1999, 14) suggest that the accurate analysts' tendencies to ignore certain information "may be a function of its relevance, complexity, or both," but they do not test this conjecture. Finally, Chen and Schoderbek (2000) find that analysts failed to incorporate the deferred tax effect of an increase in the corporate tax rate from 34 percent to 35 percent stemming from the Omnibus Budget Reconciliation Act of 1993 into their forecasts of earnings for the third quarter of 1993. However, analysts appropriately revised their forecasts of 1993 annual earnings. In a follow-up study, Chen et al. (2002) report that some analysts *inappropriately* revised their earnings forecasts for the fourth quarter of 1993 and the first quarter of 1994 to reflect a continuing effect of this one-time change in tax rates. Chen et al. (2002, Abstract) suggest that analysts' incorrect responses to the increase in tax rates may arise because of "the complexity of this deferred tax adjustment." They do not measure the relative complexity of the adjustments or other transitory items to test this conjecture, however.

In summary, judgment/decision-making research consistently finds that higher task complexity leads decision makers to select analytically simpler strategies to complete a task, and analyst-related research finds that analysts fail to use all available information in forming their forecasts. If less complex information is less costly to use than more complex information (i.e., it takes less time, effort, or training) and there are constraints on an analyst's time, effort, or ability, then, all else equal, the net benefit of incorporating information into a forecast should be a decreasing function of the information's complexity.

Based on this prior research I expect the following. If analysts impound the information in *less* complex tax-law changes into their forecasts, then analysts' *revisions*, or changes in their ETR forecasts will incorporate their estimates of the effects of the less complex tax-law changes on firms' ETRs. As a result, the new *levels* of their ETR forecasts should fully impound the estimated effects of these tax-law changes on the firms' ETRs, and thus the *error* in the level of analysts' ETR forecasts will be unrelated to the estimated effect of these less complex tax-law changes. Conversely, if analysts fail to incorporate the information in the *more* complex tax-law changes in *revising* their ETR forecasts (either because the more complex information affords constrained analysts a smaller net benefit or because analysts are less proficient in using more complex information), then the magnitude of the *error* in analysts' ETR forecasts will be related to the estimated effects of these complex tax-law changes, but their *revisions* of ETR forecasts will not. These expectations are formally stated below.

---

[4] I consider information complexity as one facet of task complexity.

[5] Maines and McDaniel (2000) find that the reporting format of comprehensive income affects *nonprofessional* investors' weighting, rather than acquisition, of the information.

H_1: The less complex the economic effects of the tax-law change on the firm, the more likely analysts' ETR forecast *revisions* reflect the effects of the tax-law change.

H_2: The more complex the economic effects of the tax-law change on the firm, the more likely the magnitude of analysts' ETR forecast *errors* will be associated with the effects of the tax-law change.

My research design isolates when analysts appear to impound information into their revisions of ETR forecasts (H_1) and when they do not (H_2). I infer that information complexity adversely affects analysts' use of that information if the less complex information explains forecast revisions and the more complex information explains forecast errors.

III. RESEARCH DESIGN AND SAMPLE

Regression Models

Testing H_1: Model 1

I test H_1 by regressing analysts' ETR forecast revisions (which capture changes in analysts' expectations due to the tax-law changes) on proxies for the effects of six tax-law changes enacted by the Tax Reform Act of 1986:

$$\begin{aligned} FR_i = {} & \alpha + \beta_{R1}STATUTORY_i + \beta_{R2}CAPITALGAIN_i + \beta_{R3}ITCREDIT_i \\ & + \beta_{R4}R\&DCREDIT_i + \beta_{R5}FOREIGNCREDIT_i + \beta_{R6}ALTERNATIVE_i \\ & + \delta_1 FRPRETAX_i + \delta_2 INTERACTION_i + \varepsilon_i \end{aligned}$$

where, for firm i:

FR_i = Value Line forecast of firm i's 1987 ETR made in the fourth quarter of 1986, less the Value Line forecast of firm i's 1987 ETR made in the third quarter of 1986;

$STATUTORY_i$ = 46 percent (the previous corporate tax rate) less the blended tax rate (based on firm i's fiscal year end—varies from 45 percent to 34 percent);

$CAPITALGAIN_i$ = mean percentage effect of the capital gains tax rate on firm i's ETR reported for 1983–1985;

$ITCREDIT_i$ = mean percentage effect of the investment tax credit on firm i's ETR reported for 1983–1985;

$R\&DCREDIT_i$ = mean percentage effect of the R&D credit on firm i's ETR reported for 1983–1985;

$FOREIGNCREDIT_i$ = mean percentage effect of the foreign tax credit on firm i's ETR reported for 1983–1985;

$ALTERNATIVE_i$ = 1 if firm i reported a net operating loss carryover for the fiscal year ended prior to September 1986 in excess of 90 percent of the firm's average pretax income over the prior three years, 0 otherwise;

$FRPRETAX_i$ = Value Line forecast of firm i's 1987 pretax income made in the fourth quarter of 1986, less the Value Line forecast of firm i's 1987 pretax income made in the third quarter of 1986, scaled by absolute value of the Value Line forecast of firm i's 1987 pretax income forecast made in the fourth quarter of 1986; and

$INTERACTION_i$ = $STATUTORY_i$ × mean percentage effect of the foreign tax rate differential on firm i's ETR reported for 1983–1985.

To form proxies for the effects of the tax-law changes, I use three types of measures. First, to proxy for the change in the statutory tax rate (STATUTORY), I use the explicit difference between the old and new statutory rate, which varies across firms as a function of their fiscal year-ends. Second, to proxy for the expected effects of changes to the capital gains tax rate (CAPITALGAIN), the investment tax credit (ITCREDIT), the research and development credit (R&DCREDIT), and the foreign tax credit (FOREIGNCREDIT), I use the mean of the percentage effect on ETR of the related tax law that each firm reported in its tax footnote over the three-year period prior to the enactment of the Tax Reform Act of 1986 (i.e., the three fiscal years ended prior to September 30, 1986). This period yields estimates of the effects of these tax laws that do not reflect the effect of changes enacted in the Tax Reform Act of 1986.[6]

Third, a firm is generally subject to the alternative minimum tax if it reduces its taxable income to less than 10 percent of its current year income by using net operating loss carryovers. Thus, I estimate the effect of the alternative minimum tax-law change on firm i's ETR using an indicator variable (ALTERNATIVE) that equals 1 if, for the fiscal year ended prior to September 1986, the firm reports a net operating loss carryover in excess of 90 percent of the average pretax book income over the prior three years, and 0 otherwise.[7]

The regression includes two control variables to capture possible sources of cross-sectional differences in ETR forecast revisions that may be correlated with my variables of interest: (1) forecast revisions of pretax income scaled by the absolute value of the fourth-quarter forecast of pretax income (FRPRETAX),[8] and (2) STATUTORY × the mean percentage effect of the foreign tax rate differential on ETR reported for 1983–1985 (INTERACTION). I include FRPRETAX as a control variable for two reasons. First, since ETRs are a function of pretax income (Wilkie and Limberg 1992), FRPRETAX controls for the cross-sectional variation in ETR forecast revisions due to anticipated changes in pretax income that are unrelated to taxes. Second, as prior research documents that firms responded strategically to tax-law changes that affect both pretax income and ETRs (e.g., Gramlich 1991; Boynton et al. 1992; Collins and Shackelford 1992; Dhaliwal and Wang 1992), I include FRPRETAX as a control for the cross-sectional variation in ETR forecast revisions related to these anticipated strategic responses.

I control for STATUTORY × the mean percentage effect of the foreign tax rate differential reported for 1983–1985 (i.e., INTERACTION), because firms with foreign taxable income will not recognize the full effect of the change in the statutory rate, due to reported foreign taxable income.

---

[6] As an alternative proxy for the effect of these four tax laws on firms' ETRs, I could have used the difference between the actual amount of the 1987 reconciling item and the average effect of the reconciling item in past years. However, I chose to limit the proxies to information included in firms' footnotes prior to the forecast date to ensure that the information was publicly available to the analysts at the time they made their ETR forecasts.

[7] Omer and Atwood (1991) report that two-thirds of firms paying the alternative minimum tax did so because the tax law limited their use of net operating loss carryovers when calculating the alternative taxable income. In addition, of the six variables included in their study, Manzon and Plesko (1999) report that NOL is the best indicator of firms that reported paying the alternative minimum tax.

[8] Because Value Line does not explicitly forecast pretax income, I calculate it by dividing forecasted net profit by (1 − forecasted ETR). For consistency with the dependent variable (ETR), I deflate pretax income forecast revisions by the absolute value of the 1987 pretax income forecast made in the fourth quarter of 1986.

Prior research documents that analyst-specific attributes, such as experience, employer size, and the number of firms and industries an analyst follows, are associated with analyst accuracy (Mikhail et al. 1997; Clement 1999). Value Line's employment of all the analysts in my sample controls for variation in analyst accuracy associated with employer size (Clement 1999). Although I am unable to control for differences in other analyst attributes, such as firm-specific forecasting experience, these other attributes generally have a modest effect on analyst accuracy (Jacob 1997), so this limitation is unlikely to affect my inferences.

I expect five of the six tax-law changes to have unambiguous directional effects on firm ETRs, allowing for one-tailed tests of significance. Reductions in the statutory rate and the extension of the R&D credit will reduce ETRs; therefore, I expect the coefficients on STATUTORY and R&DCREDIT to be negative. Eliminating the capital gains tax rate and the investment tax credit and limiting the use of the foreign tax credit will increase ETRs; therefore I expect the coefficients on CAPITALGAIN, ITCREDIT, and FOREIGNCREDIT to be positive.

The net effect of changes to the alternative minimum tax across firms is ambiguous. Many of the "tax adjustment" items included in the calculation of the alternative minimum tax are temporary differences, which will generate an alternative minimum tax credit for use in future years. For these items, the alternative minimum tax decreases expected ETR because previously no alternative minimum tax credit was generated.[9] Conversely, "preference items" used to calculate the alternative minimum tax that arise due to permanent differences do not generate an alternative minimum tax credit. For these items, the alternative minimum tax increases future ETR. Finally, the limitation applied to the alternative minimum tax net operating loss may effect whether a firm eventually uses a net operating loss, which changes the net operating loss from a temporary difference to a permanent difference. As I am unable to form an unambiguous expectation for the ALTERNATIVE variable, the test of significance of its coefficient in the regression is two-tailed.

Testing H_2: Model 2

If analysts fully incorporate less complex information into their forecasts of ETRs, but they do not fully incorporate more complex information, then the magnitude of the ETR *forecast errors* will (1) not be associated with the less complex information, but (2) will be associated with the more complex information. To test this hypothesis, I regress the absolute value of the analysts' ETR forecast errors on the proxies for the effects of the six tax-law changes and two control variables:

$$\begin{aligned} |FE_i| = {} & \alpha_i + \beta_{R1}\mathrm{STATUTORY}_i + \beta_{R2}\mathrm{CAPITALGAIN}_i + \beta_{R3}\mathrm{ITCREDIT}_i \\ & + \beta_{R4}\mathrm{R\&DCREDIT}_i + \beta_{R5}\mathrm{FOREIGNCREDIT}_i + \beta_{R6}\mathrm{ALTERNATIVE}_i \\ & + \beta_1|\mathrm{FEPRETAX}_i| + \beta_2\mathrm{INTERACTION}_i + \varepsilon \end{aligned}$$

---

[9] During my sample period, corporate alternative minimum tax credits arose only for alternative minimum tax paid due to temporary differences. The Revenue Reconciliation Act of 1989 changed this rule such that, for tax years beginning after 1989, corporations received alternative minimum tax credits for the entire amount of their alternative minimum tax paid.

where, for firm i:

$|FE_i|$ = absolute value of the error in the Value Line forecast of firm i's 1987 ETR—the actual 1987 ETR less the Value Line forecast of firm i's 1987 ETR forecast made in the fourth quarter of 1986; and

$|FEPRETAX_i|$ = the absolute value of the ratio of the error in the Value Line forecast of firm i's 1987 pretax income, scaled by the actual 1987 pretax income. The error in the Value Line forecast of firm i's 1987 pretax income is the actual 1987 pretax income less the Value Line forecast of firm i's 1987 pretax income forecast made in the fourth quarter of 1986.

All other variables are as defined above.

The dependent variable is the absolute value of the ETR forecast errors. I expect the effects of tax-law changes that are not fully incorporated into analysts' forecasts of ETRs to be positively associated with the magnitude of the errors in analysts' ETR forecasts, allowing for one-tailed significance tests.

Complexity Rankings of the Six Tax-Law Changes

For each of the six tax-law changes, three certified public accountants practicing during the enactment of the Tax Reform Act of 1986 completed the 15 questions that comprise the AICPA's tax-complexity index.[10] For each question, participants responded on a scale ranging from −3 (an expected decrease in complexity) to +3 (an expected increase in complexity). The sum of the scores from the 15 questions measures the complexity of each tax-law change, for each participant. I calculate the mean of these sums across the three participants and use them to rank the tax-law changes by complexity.[11]

I assess the consistency of the rankings obtained across the participants using Kendall's coefficient of concordance, a measure of consistency of rankings across more than two individuals. Kendall's W is 0.91, indicating high consensus.[12] The only inconsistency in complexity ranking across the individuals is a disagreement, in one case, about whether the change in the capital gains tax rate or the change in investment tax credit was more complex.

This process produces a *relative* measure of complexity, not an *absolute* measure. Table 1 shows that three of the tax-law changes score as relatively simple: change in statutory rate = 1.7; change in capital gains rate = 2.3; and change in investment tax credit = 2.7. For the first two items, the Tax Reform Act of 1986 changed the rate applied to taxable income; for the third item, the Act effectively eliminated the tax credit. In contrast, the remaining three tax-law changes score as much more complex: change in R&D credit = 20.3; change in foreign tax credit = 27.0; change in alternative minimum tax = 31.0. The Act extended the R&D credit, but at a reduced rate of 20 percent with stricter guidelines about what qualifies as research. The Act substantially changed a firm's ability to use foreign tax credits by requiring it to compute the foreign tax credit and apply the credit

---

[10] The AICPA designed the tax-complexity index for accountants and members of Congress to assess whether a proposed tax-law change would increase or decrease the overall complexity of the tax system. Generally, the questions in the index relate to definitions or concepts that may change due to the tax-law change, changes or additions to taxpayers' calculations, when the tax-law change would be implemented, and taxpayer recordkeeping requirements.

[11] Using median values does not affect the complexity rankings.

[12] If the rankings across all the individuals are exactly the same (no disagreement in relative complexity), Kendall's W will be 1. If there is complete disagreement across all individuals, then Kendall's W will be 0 or very close to 0 (Conover 1980, 305).

TABLE 1
Complexity Index Scores for Six Tax-Law Changes Enacted by the Tax Reform Act of 1986[a]

Tax-Law Change	*Variable Name*	*Complexity Score*[b]
Statutory tax rate	STATUTORY	1.7
Capital gains tax rate	CAPITALGAIN	2.3
Investment tax credit	ITCREDIT	2.7
Research and development credit	R&DCREDIT	20.3
Foreign tax credit	FOREIGNCREDIT	27.0
Alternative minimum tax	ALTERNATIVE	31.0

[a] The tax-law changes: decrease in the statutory tax rate (STATUTORY); increase in the capital gains tax rate (CAPITALGAIN); elimination of the investment tax credit (ITCREDIT); changes in the computation of the available R&D tax credit (R&DCREDIT); extensive changes to the foreign tax credit (FOREIGNCREDIT); and implementation of an alternative minimum tax (ALTERNATIVE).

[b] The complexity scores are based on responses from three certified public accountants practicing during the enactment of the Tax Reform Act of 1986 to 15 questions that comprise a tax-complexity index designed by AICPA.

limitations separately for different categories of income. The Act also replaced the minimum tax with an alternative minimum tax system. Overall, this complexity ranking appears appropriate, given the tax-law changes made by the Tax Reform Act of 1986.

In the context of the tax laws examined in this study, *information complexity* is a function of both rule complexity (understanding and applying a tax law to a static set of facts) and strategic complexity (understanding firms' tax-planning responses to new tax laws). One limitation of this study is that my empirical measure of complexity ranks tax-law changes in terms of their rule complexity only. The validity of my tests rests on whether this ranking reflects the overall information complexity of these tax laws. If the unmeasured strategic complexity would alter the rankings based on the rule complexity, then the conclusions drawn may be incorrect. I mitigate this measurement error problem by controlling for pretax income to capture firms' tax-planning responses to the Tax Reform Act of 1986 that affected pretax accounting income and ETRs. In addition, in my discussion of the results, I evaluate anecdotal evidence of firm strategic responses documented in prior studies and the likely effects of these strategic responses on the complexity rankings.

Sample

My sample consists of 355 firms that Value Line followed from 1984 through 1988. Value Line provides both forecasts and reported actual amounts of ETRs and earnings.[13] It does not provide ETR forecasts for firms with either negative or zero taxes or earnings; thus my analysis includes only firms Value Line expected to be profitable in the near future.

Value Line forecasts annual ETRs and earnings once every 13 weeks (i.e., once each calendar quarter) for each firm, issuing annual forecasts for the current year, the subsequent year, and three to five years ahead. Since Congress enacted the Tax Reform Act of 1986 in October of that year, I consider Value Line forecasts made in the third quarter of 1986 (July–September) as prior to enactment and forecasts made in the fourth quarter of 1986 (October–December) as after enactment.[14] I use Value Line ETR forecasts for the fiscal years ended May 31, 1987, through April 30, 1988 (the "subsequent year"), to ensure that my tests use forecasts for the year following the tax-law changes.[15]

---

[13] For 44 percent of the firms, Value Line's "actual" firms' ETRs differ from ETRs calculated from Compustat. As Value Line ETRs are the appropriate benchmark for assessing Value Line analysts' forecast accuracy, I hand-collect both the actual and forecast data from Value Line.

Value Line issued forecasts of ETRs for 1,225 firms in the third and fourth quarters of 1986. Additional data (Value Line reported "actual" ETRs for 1984, 1985, and 1987 and Value Line forecasted ETRs for 1985 and 1986 issued in the fourth quarter of 1983 and 1984) required to form the forecast-error measures in the primary analysis and the robustness tests limit my sample to 424 firms. I found the firm-specific tax footnote data from annual reports (required to estimate the proxies for the effects of the tax-law changes) for 360 of these firms.[16] I eliminated five firms considered influential in my regressions based on the Belsley et al. (1980, 28) procedure. Thus, the final sample includes 355 firms.[17]

Table 2 provides descriptive statistics for this sample. As shown in Panel A, the mean and median ETR forecast revisions are negative (−1.48 and −2.0, respectively), consistent with analysts, on average, expecting ETRs to decrease around 1.5 percentage points after the implementation of the Tax Reform Act of 1986. (A firm with a forecasted ETR of 42 percent prior to the Act and a forecasted ETR of 40 percent after the Act would have a forecast revision of −2.0.) The mean (median) absolute value of the Value Line ETR forecast error is 3.63 (2.2), which indicates that, on average, analysts missed firms' ETRs by 3.6 percentage points. (A firm with a forecasted ETR of 40 percent after the Act and an actual ETR of either 36.4 percent or 43.6 percent would have an absolute ETR forecast error of 3.6.)

Panel B of Table 2 provides descriptive statistics for the proxies for the effects of the six tax-law changes. The mean values of all the tax-law proxies are greater than zero. In contrast, median values equal zero for all but STATUTORY and ITCREDIT, indicating that less than half of my sample firms are affected by the other four tax-law changes. Specifically, STATUTORY and ITCREDIT are nonzero for 99.7 percent and 85.6 percent of my sample firms, respectively, but CAPITALGAIN, R&DCREDIT, FOREIGNCREDIT, and ALTERNATIVE are nonzero for just 18.6 percent, 14.9 percent, 12.4 percent, and 10.4 percent of my sample firms, respectively. Some firms are affected only by the change in the statutory rate, while others are affected by multiple tax-law changes.

Panel C of Table 2 provides descriptive statistics for the control variables. The mean (median) pretax income forecast revision is −0.08 (−0.03); it is zero for 5.4 percent of the firms. (A value of −0.08 represents an expected 8.0 percent decrease in pretax income.) The mean (median) *absolute value* of the pretax income forecast error is 0.29 (0.15), indicating an average forecast error for pretax income of 29 percent. The mean (median)

---

[14] While the Tax Reform Act of 1986 and many of its related law changes were widely publicized prior to this time, these pre- and post-Tax Reform Act of 1986 periods are consistent with Value Line's statements at the beginning of the fourth-quarter forecasts: "As noted on the 'pink sheet' accompanying this week's Edition, Value Line will henceforth include the effects of tax-law changes in estimates of future earnings" (Value Line 1986b, 99).

[15] For example, for a Value Line forecast issued in October 1986 for a firm with a November fiscal year-end, I collected Value Line's forecast of the annual effective tax rate for the fiscal year ending November 1987.

[16] I was able to locate 364 firms with necessary tax footnote data to estimate the independent variables that proxy for the effects of the four tax-law changes. These proxies are typically positive, although, for four firms the ITCREDIT proxy was negative due to investment tax recapture. I eliminated these four firms because the relation between these negative values and the dependent measures may be different than the relation between the positive values and the dependent measures. Including these firms does not affect the study's inferences.

[17] Retaining these observations does not affect the study's inferences.

TABLE 2
Descriptive Statistics for Dependent, Explanatory, and Control Variables
n = 355 Firms

Panel A: Dependent Variables

Variable[a]	Mean	Median	Nonzero Observations[b] Percent	Nonzero Observations[b] Frequency
FR	−1.48	−2.0	87.3	310
|FE|	3.63	2.2	99.4	353

Panel B: Explanatory Variables

Variable	Mean	Median	Nonzero Observations[b] Percent	Nonzero Observations[b] Frequency
STATUTORY	5.96	6.00	99.7	354
CAPITALGAIN	0.33	0.00	18.6	66
ITCREDIT	1.92	1.70	85.6	304
R&DCREDIT	0.28	0.00	14.9	53
FOREIGNCREDIT	0.14	0.00	12.4	44
ALTERNATIVE	0.10	0.00	10.4	37

Panel C: Control Variables

Variable	Mean	Median	Nonzero Observations[b] Percent	Nonzero Observations[b] Frequency
FRPRETAX	−0.08	−0.03	94.6	336
|FEPRETAX|	0.29	0.15	100.0	355
INTERACTION	−0.07	0.00	18.1	65

---

[a] Variable definitions:

FR_i = Value Line forecast of firm i's 1987 ETR made in the fourth quarter of 1986, less the Value Line forecast of firm i's 1987 ETR made in the third quarter of 1986;

$|FE_i|$ = absolute value of the error in the Value Line forecast of firm i's 1987 ETR—the actual 1987 ETR less the Value Line forecast of firm i's 1987 ETR forecast made in the fourth quarter of 1986;

$STATUTORY_i$ = 46 percent (the previous corporate tax rate) less the blended tax rate (based on firm i's fiscal year end—varies from 45 percent to 34 percent);

$CAPITALGAIN_i$ = mean percentage effect of the capital gains tax rate on firm i's ETR reported for 1983–1985;

$ITCREDIT_i$ = mean percentage effect of the investment tax credit on firm i's ETR reported for 1983–1985;

$R\&DCREDIT_i$ = mean percentage effect of the R&D credit on firm i's ETR reported for 1983–1985;

$FOREIGNCREDIT_i$ = mean percentage effect of the foreign tax credit on firm i's ETR reported for 1983–1985;

$ALTERNATIVE_i$ = 1 if firm i reported a net operating loss carryover for the fiscal year ended prior to September 1986 in excess of 90 percent of the firm's average pretax income over the prior three years, 0 otherwise;

(continued on next page)

TABLE 2 (continued)

FRPRETAX_i = Value Line forecast of firm i's 1987 pretax income made in fourth quarter of 1986, less the Value Line forecast of firm i's 1987 pretax income made in the third quarter of 1986, scaled by absolute value of the Value Line forecast of firm i's 1987 pretax income forecast made in the fourth quarter of 1986;

$|\text{FEPRETAX}_i|$ = the absolute value of the ratio of the 1987 error in Value Line's forecast of firm i's pretax income, scaled by the actual 1987 pretax income. The 1987 pretax income forecast error is the actual 1987 pretax income less the 1987 pretax income Value Line made in the fourth quarter of 1986; and

INTERACTION_i = STATUTORY × mean percentage effect on firm i's ETR reported for 1983–1985.

[b] Percentage (number) of firms for which the variable is nonzero. For example, CAPITALGAIN is nonzero for 18.6 percent of the sample (66 firms) and zero for 81.4 percent of the sample (289 firms).

interaction term, which captures the mitigated effect of the change in statutory rates for firms with foreign operations, is −0.07 (0.00). The mean value for the interaction term suggests that, for the 65 firms that reported a foreign tax rate differential, that differential was negative.

To assess whether analysts exercised due care in making their ETR forecasts, I examine (but do not report) the mean and median values of pretax income and ETR forecast errors. Consistent with prior research on forecasts of *earnings* forecast errors (e.g., O'Brien 1988; Philbrick and Ricks 1991), I find that mean pretax income forecasts are significantly optimistic, but median pretax income forecasts are, on average, unbiased. The mean and median ETR forecast errors (0.13 and 0.20, respectively) are not significantly different from zero, documenting that analysts made unbiased ETR forecasts for 1987. ETR forecast errors range from 2.4 at the 75th percentile to −2.2 at the 25th percentile.

IV. EMPIRICAL RESULTS

Test of H_1: Value Line Analysts' Revisions of Their ETR Forecasts

I test H_1 by estimating Model 1, which provides evidence on the extent to which analysts' ETR forecast revisions are associated with the effects of the tax-law changes. Table 3, Panel A lists the coefficients for each of the six tax laws in order of their relative complexity (R1 [least complex] to R6 [most complex]).

The overall regression is highly significant with an adjusted R^2 of 24.6 percent, indicating that the proxies capture cross-sectional variation in the forecast revisions. The results suggest that Value Line analysts revised their forecasts of firms' ETRs in response to the three tax-law changes ranked as least complex. As expected, analysts reduced their forecasts of ETRs in response to the expected reduction in the firm's statutory tax rate (STATUTORY), and they increased their forecasts of ETRs in response to the expected increase in the firm's capital gains tax rate (CAPITALGAIN) and the effect on the firm of the virtual elimination of the investment tax credit (ITCREDIT) (all at p-values of 0.001). In contrast, the revisions of ETR forecasts are unrelated to the three tax-law changes ranked as most complex; R&DCREDIT, FOREIGNCREDIT, and ALTERNATIVE provide no explanatory value in the forecast revision regression. With respect to the control variables, the positive coefficient on FRPRETAX (at a p-value of 0.001), suggests that upward revisions in forecasts of pretax income are associated with upward revisions in ETR forecasts. The coefficient on INTERACTION is not significant.[18]

---

[18] I also estimate the regressions including the main effect of the foreign tax rate differential in addition to the INTERACTION term. The main effect variable of the foreign tax rate differential is never significant and the overall results are substantively similar, so I do not include the main effect variable in the reported regressions.

TABLE 3

Regressions of Revisions (Panel A) and Absolute Errors (Panel B) in Value Line Forecasts of Effective Tax Rates on Proxies for the Effects of Six Tax-Law Changes of Increasing Complexity Enacted by the Tax Reform Act of 1986

Panel A: Forecast Revisions[a]

$$FR_i = \alpha_i + \beta_{R1}STATUTORY_i + \beta_{R2}CAPITALGAIN_i + \beta_{R3}ITCREDIT_i + \beta_{R4}R\&DCREDIT_i + \beta_{R5}FOREIGNCREDIT_i + \beta_{R6}ALTERNATIVE_i + \delta_1 FRPRETAX_i + \delta_2 INTERACTION_i + \varepsilon$$

		Complexity Rank[a]								
Variable[b]	*Intercept*	β_{R1}	β_{R2}	β_{R3}	β_{R4}	β_{R5}	β_{R6}	δ_1	δ_2	*Adj.* R^2
Predictions	?	−	+	+	−	+	?	?	?	
Coefficient	2.39	−0.68	0.38	0.23	−0.15	−0.10	−0.61	3.23	0.02	24.6%
t-statistic	2.43	−4.28***	3.21***	5.92***	−1.33	−0.68	−0.95	5.17###	1.55	

Panel B: Absolute Forecast Errors[a]

$$|FE_i| = \alpha_i + \beta_{R1}STATUTORY_i + \beta_{R2}CAPITALGAIN_i + \beta_{R3}ITCREDIT_i + \beta_{R4}R\&DCREDIT_i + \beta_{R5}FOREIGNCREDIT_i + \beta_{R6}ALTERNATIVE_i + \delta_1|FEPRETAX_i| + \delta_2 INTERACTION_i + \varepsilon$$

		Complexity Rank[a]								
Variable[b]	*Intercept*	β_{R1}	β_{R2}	β_{R3}	β_{R4}	β_{R5}	β_{R6}	δ_1	δ_2	*Adj.* R^2
Coefficient	4.11	−0.27	−0.15	0.07	0.21	0.73	1.96	2.84	0.01	15.0%
t-statistic	3.67	−1.39	−1.07	1.47	1.36	4.24***	2.53**	7.38***	0.05	

(*continued on next page*)

TABLE 3 (continued)

---

*, **, *** $p < 0.05$, $p < 0.01$, $p < 0.001$, respectively, one-tailed; #, ##, ### $p < 0.05$, $p < 0.01$, $p < 0.001$, respectively, two-tailed.

[a] The six tax-law variables are ordered from least complex (R1) to most complex (R6).

[b] Variable definitions:

FR_i = Value Line forecast of firm i's 1987 ETR made in the fourth quarter of 1986, less the Value Line forecast of firm i's 1987 ETR made in the third quarter of 1986;

$|FE_i|$ = absolute value of the error in the Value Line forecast of firm i's 1987 ETR—the actual 1987 ETR less the Value Line forecast of firm i's 1987 ETR forecast made in the fourth quarter of 1986;

$STATUTORY_i$ = 46 percent (the previous corporate tax rate) less the blended tax rate (based on firm i's fiscal year end—varies from 45 percent to 34 percent);

$CAPITALGAIN_i$ = mean percentage effect of the capital gains tax rate on firm i's ETR reported for 1983–1985;

$ITCREDIT_i$ = mean percentage effect of the investment tax credit on firm i's ETR reported for 1983–1985;

$R\&DCREDIT_i$ = mean percentage effect of the R&D credit on firm i's ETR reported for 1983–1985;

$FOREIGNCREDIT_i$ = mean percentage effect of the foreign tax credit on firm i's ETR reported for 1983–1985;

$ALTERNATIVE_i$ = 1 if firm i reported a net operating loss carryover for the fiscal year ended prior to September 1986 in excess of 90 percent of the firm's average pretax income over the prior three years, 0 otherwise;

$FRPRETAX_i$ = Value Line forecast of firm i's 1987 pretax income made in fourth quarter of 1986, less the Value Line forecast of firm i's 1987 pretax income made in the third quarter of 1986, scaled by absolute value of the Value Line forecast of firm i's 1987 pretax income forecast made in the fourth quarter of 1986;

$|FEPRETAX_i|$ = the absolute value of the ratio of the error in the Value Line forecast of firm i's 1987 pretax income, scaled by the actual 1987 pretax income. The error in the Value Line forecast of firm i's 1987 pretax income is the actual 1987 pretax income less the Value Line forecast of firm i's 1987 pretax income forecast made in the fourth quarter of 1986; and

$INTERACTION_i$ = STATUTORY × mean percentage effect on firm i's ETR reported for 1983–1985.

The evidence presented in Panel A of Table 3 supports my hypothesis that analysts use less complex information to a greater extent than they use more complex information. Two alternative interpretations of these results, however, are that analysts' forecast revisions did not impound the more complex information because it does not help them accurately forecast firms' ETRs, or because the proxy for the effect of the tax-law change is not well specified. Testing H_2 allows me to distinguish between my hypothesis and these two alternative explanations.

Test of H_2: Absolute Value of Value Line Analysts' ETR Forecast Errors

Panel B of Table 3 reports the results of my test of H_2, where I regress the absolute value of the Value Line analysts' ETR forecast errors on proxies for the effects of the six tax-law changes and the control variables. The overall regression is significant with an adjusted R^2 of 15.0 percent, evidence that the proxies capture cross-sectional variation in the absolute forecast errors. Consistent with H_2, analysts do not appear to impound the effects of the most complex tax-law changes on firms' ETRs. Specifically, the two most complex tax-law changes (FOREIGNCREDIT and ALTERNATIVE) are significantly positively associated with the magnitude of analysts' ETR forecast errors (at p-values of 0.001 and 0.01, respectively). In contrast, the three least complex tax-law changes, that are significantly associated with analysts' forecast *revisions*, are unrelated to the magnitude of the error in analysts' forecasts of ETRs, as if the analysts have fully assimilated the implications of these three tax-law changes for firms' ETRs. The proxy for the effect of the changes to the R&D tax credit (R&DCREDIT) is not significantly associated with either the ETR forecast revisions or the absolute ETR forecast errors, suggesting that either the proxy may not be well specified or is not relevant. Finally, the significantly positive coefficient on the control variable, |FEPRETAX|, indicates that larger errors in analysts' pretax income forecasts are associated with larger errors in their ETR forecasts.

The determinants of forecast revisions and forecast errors together provide strong evidence that complexity adversely affects Value Line analysts' use of tax information. The effects of the three least complex tax-law changes explain much of the variation in analysts' *revisions* of their ETR forecasts, but are unrelated to the magnitude of the *errors* in analysts' ETR forecasts, which suggests the analysts correctly used the information to forecast ETRs. In contrast, the two tax-law changes ranked as most complex are associated with variation in the magnitude of analysts' ETR forecast *errors*, but are not associated with analysts' *revisions* of their ETR forecasts. These results suggest that higher information complexity reduces analysts' use of that information.

Rule Complexity vs. Strategic Complexity

As discussed earlier, in the context of the tax-law changes, overall information complexity is a function of both rule complexity and strategic complexity, but my empirical measure ranks tax-law changes only in terms of rule complexity. The validity of my inferences rests on the assumption that the ranking based on the measured rule complexity reflects the true (unobservable) ranking based on overall complexity. Evidence from prior research supports this assumption.

Prior research documents that firms responded strategically to three of the tax-law changes included in my study: changes in the statutory rate, changes in the foreign tax credit, and changes in the alternative minimum tax. Several studies (e.g., Manzon 1992; Scholes et al. 1992; Guenther 1994; Maydew 1997) document that firms shifted taxable income or deductions across time in response to decreases in the statutory tax rate (a relatively low-complexity tax-law change). Because this strategic response affected taxable and financial income equally, ETRs are unaffected. The strategic complexity of this firm response seems to be low and is uncorrelated with ETRs.

Collins and Shackelford (1992) document that firms issued preferred stock in lieu of debt in response to the Tax Reform Act of 1986's changes to the foreign tax credit, a more complex tax-law change. By doing so, firms increased financial income but decreased tax expense by increased foreign tax-credit utilization, a strategic response that would reduce firms' ETRs and appears relatively complex. Several studies (e.g., Gramlich 1991; Boynton et al. 1992; Dhaliwal and Wang 1992) examine other seemingly complex strategic responses to changes made to the alternative minimum tax, the tax-law change my study ranked as most complex. Firms' strategic responses typically affected the calculation of the alternative minimum taxable income, which is taxed at a different rate than regular taxable income (20 percent vs. 34 percent), resulting in a net effect on firms' ETRs. The strategic complexity of the firms' responses related to both foreign tax credits and the alternative minimum tax, which may involve issuing preferred stock or changing investment decisions or accounting policies, seems to be high, which would increase the overall complexity of these tax-law changes

In summary, the tax-law change ranked as lowest in complexity elicited a straightforward strategic response that would not affect ETRs, while two of the tax-law changes ranked as highest in complexity elicited more complex strategic responses that would affect ETRs. These results suggest that unmeasured strategic complexity would accentuate, rather than alter, the complexity rankings of the six-tax law changes.

Robustness Test

I use the magnitude of ETR forecast errors as the dependent variable in Model 2. However, a firm's 1987 ETR forecast error might be related to the underlying complexity of applying the old tax laws, which varies across firms. To control for this potential confounding factor, I reestimate Model 2 using the *change* in the absolute value of the ETR forecast error as the dependent measure, calculated as follows. I first calculate the absolute forecast errors for 1987 as in Table 3, and I compute analogous absolute forecast errors for 1984 and 1985 (as the absolute value of the year t actual ETR less the Value Line year t ETR forecast made in the fourth quarter of year t−1). I then calculate the difference between (1) the absolute 1987 ETR forecast error and (2) the average of the absolute 1984 and 1985 ETR forecast errors. This difference is the change in the magnitude of the ETR forecast errors—the magnitude of the ETR forecast error less the pre-Tax Reform Act of 1986 average absolute forecast error.[19]

Table 4 reports the results from the respecified regression. The estimated effects of R&DCREDIT, FOREIGNCREDIT, and ALTERNATIVE are associated with increases in the magnitude of analysts' errors in forecasts of ETRs (at p-values of 0.01, 0.01, and 0.05, respectively). Both control variables, change in the absolute pretax income forecast error and INTERACTION, are associated with increases in the magnitude of analysts' errors in forecasting ETRs after the Tax Reform Act of 1986 (at p-values of 0.001 and 0.05, respectively). Although these results are not identical to those documented in Table 3, the relation between complex information and the magnitude of analysts' ETR forecast errors is the same; more complex information items explain increases in absolute forecast errors of ETRs, whereas less complex information items do not.

---

[19] To be consistent, I use the change in the absolute value of pretax income forecast error in this specification of Model 2.

TABLE 4

Regression of Change in Absolute Errors in Value Line Forecasts of Effective Tax Rates on Proxies for the Effects of Six Tax-Law Changes of Increasing Complexity Enacted by the Tax Reform Act of 1986—Alternative Specification[a]

$$\Delta|FE| = \alpha + \beta_{R1}STATUTORY_i + \beta_{R2}CAPITALGAIN_i + \beta_{R3}ITCREDIT_i + \beta_{R4}R\&DCREDIT_i + \beta_{R5}FOREIGNCREDIT_i + \beta_{R6}ALTERNATIVE_i + \delta_1\Delta|FEPRETAX_i| + \delta_2INTERACTION_i + \varepsilon$$

Variable[b]	*Intercept*	Complexity Rank[a]								
		β_{R1}	β_{R2}	β_{R3}	β_{R4}	β_{R5}	β_{R6}	δ_1	δ_2	*Adj.* R^2
Coefficient	5.64	−0.39	0.02	0.06	0.43	0.68	1.50	3.67	0.50	16.7%
t-statistic	3.53	−1.54	0.52	1.52	2.37**	2.70**	1.69*	5.95###	1.97#	

---

*, **, *** $p < 0.05$, $p < 0.01$, $p < 0.001$, respectively, one-tailed; #, ##, ### $p < 0.05$, $p < 0.01$, $p < 0.001$, respectively, two-tailed.

[a] The six tax-law variables in the regression are ranked by complexity, from least complex (R1) to most complex (R6).

[b] Variable definitions:

$\Delta|FE_i|$ = difference between the absolute value of the 1987 ETR forecast error and the absolute value of the average of the firm's 1984 and 1985 ETR forecast errors. The 1987 ETR forecast error is the actual 1987 ETR less the Value Line forecast of firm i's 1987 ETR forecast made in the fourth quarter of 1986. I compute analogous absolute forecast errors for 1984 and 1985;

$STATUTORY_i$ = 46 percent (the previous corporate tax rate) less the blended tax rate (based on firm i's fiscal year end—varies from 45 percent to 34 percent);

$CAPITALGAIN_i$ = mean percentage effect of the capital gains tax rate on firm i's ETR reported for 1983–1985;

$ITCREDIT_i$ = mean percentage effect of the investment tax credit on firm i's ETR reported for 1983–1985;

$R\&DCREDIT_i$ = mean percentage effect of the R&D credit on firm i's ETR reported for 1983–1985;

$FOREIGNCREDIT_i$ = mean percentage effect of the foreign tax credit on firm i's ETR reported for 1983–1985;

$ALTERNATIVE_i$ = 1 if firm i reported a net operating loss carryover for the fiscal year ended prior to September 1986 in excess of 90 percent of the firm's average pretax income over the prior three years, 0 otherwise;

$\Delta|FEPRETAX_i|$ = the difference between the absolute value of the ratio of the error in Value Line's forecast of firm i's 1987 pretax income, scaled by the actual 1987 pretax income and the mean of the absolute value of analogous measures for 1984 and 1985. The 1987 pretax income forecast error is the actual 1987 pretax income less the Value Line forecast of firm i's 1987 pretax income made in the fourth quarter of 1986; and

$INTERACTION_i$ = $STATUTORY_i$ × mean percentage effect on firm i's ETR reported for 1983–1985.

V. CONCLUSIONS

This study provides evidence on the relation between information complexity and analysts' use of that information. I examine how six specific tax-law changes enacted by the Tax Reform Act of 1986 affect Value Line analysts' effective tax rate (ETR) forecasts for 355 firms. I show that analysts' *revisions* of their ETR forecasts appear to impound the effects of the less complex tax-law changes but not the effects of the more complex tax-law changes. As expected, if analysts impound the less complex information but not the more complex information, the effects of the less complex tax-law changes are not associated with the magnitude of analysts' errors in their ETR forecasts, but the effects of the more complex tax-law changes do explain variation in the magnitude of ETR forecast errors. In conjunction, these results are consistent with increased complexity of information reducing the quality of analysts' forecasts based on that information.

Empirical evidence that the complexity of information imposes sufficient costs on even expert users of financial information—analysts—to reduce their use of that information is relevant to the Financial Accounting Standards Board and the SEC as they consider changing what information companies are required to disclose, and when and how they must disclose it. In addition, the results suggest that researchers should consider the effect of the complexity of information in designing and interpreting the results of their research. For example, studies that find no significant relation between specific information and stock prices changes (or other firm-specific changes) without considering the complexity of the information may erroneously conclude that the information is irrelevant, when it is actually relevant but too complex for users to fully incorporate.

My results are subject to two alternative explanations. First, if the proxies for the effects of the tax-law changes are noisier for the more complex tax-law changes than for the less complex tax-law changes, then the results could reflect the differential quality of the proxies rather than the differential complexity of the tax-law changes. For example, the proxy for the least complex tax-law change is a function of a firm's fiscal year-end month—which is precise—while the proxy for the most complex tax-law change is perhaps the most noisy—it is an indicator variable. This is unlikely to provide a complete explanation for my results, however. Four of the proxies used to capture the effects of the tax-law changes are based on detailed data drawn from each firm's tax footnote (CAPITALGAIN, ITCREDIT, R&DCREDIT, and FOREIGNCREDIT). This should reduce the variance in the noise across these proxies and hence the concern that the results reflect the precision of the proxies.

The second alternative explanation for my results is that the more complex tax laws affect fewer firms, as shown in Table 2, so analysts choose not to invest their time analyzing the potential effect of those tax-law changes. However, the potential benefit of properly incorporating the two most complex tax variables appears to be material; the average estimated error in net profit due to the failure to incorporate the effects of these tax laws ranges between $1.93 and $4.61 million.[20]

---

[20] Multiplying the coefficients on FOREIGNCREDIT and ALTERNATIVE of 0.73 and 1.96, respectively (Table 3, Panel B) by their mean values of 0.14 and 0.10, respectively (Table 2, Panel B) yields an average forecast error due to under-using available information in the amount of 0.102 and 0.196, respectively. These tax-law changes affect 12.4 percent and 10.4 percent of the firms in my sample; thus, an estimate of the average error of under-using the information would be 0.82 percent and 1.96 percent, respectively. With a mean forecasted net profit of $235 million for my sample firms, the average estimated error in net profit due to the failure to incorporate the effects of these tax laws ranges between $1.93 and $4.61 million. These values must be interpreted with caution, because the calculations are based on point estimates from a regression and mean firm-specific values. However, they provide a rough estimate of the potential improvement in analysts' forecast accuracy from incorporating information related to these complex tax-law changes.

I capture the key construct in this study, information complexity, by ranking the tax-law changes by complexity; information complexity is not an independent variable in the tests. This is a limitation of my study, as the research design does not provide a direct test of the effect of complexity on analysts' information use. Instead, one must infer the effect by observing the pattern of statistical significance of the coefficients of the proxies for the tax-law changes in two related regressions. Nevertheless, the study's results illustrate the importance of considering information complexity when examining analysts' use of or failure to use information.

This study does not examine why analysts fail to incorporate more complex information into their forecasts. Either analysts' abilities to incorporate information into their forecasts correctly is a decreasing function of the complexity of the information, or analysts choose not to assimilate complex information because the cost of using the information exceeds the benefit of doing so. Distinguishing between these alternative explanations for why analysts under-use complex information may be a fruitful avenue for future research.

REFERENCES

Abarbanell, J. 1991. Do analysts' earnings forecasts incorporate information in prior stock price changes? *Journal of Accounting and Economics* 14 (2): 147–165.

———, and V. Bernard. 1992. Tests of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior. *The Journal of Finance* 47 (3): 1181–1208.

———, and B. Bushee. 1997. Fundamental analysis, future earnings, and stock price. *Journal of Accounting Research* 35 (1): 1–24.

———, and ———. 1998. Abnormal returns to a fundamental analysis strategy. *The Accounting Review* 73 (1): 19–45.

Ali, A., A. Klein, and J. Rosenfeld. 1992. Analysts' use of information about permanent and transitory earnings components in forecasting annual EPS. *The Accounting Review* 67 (1): 183–199.

American Institute of Certified Public Accountants (AICPA). 1993. *AICPA Tax Complexity Index.* New York, NY: AICPA.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York, NY: Wiley.

Bernard, V., and J. Thomas. 1989. Post-earnings announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27 (Supplement): 1–48.

———, and ———. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13 (4): 305–340.

Bettman, J., E. Johnson, and J. Payne. 1990. A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes* 45 (1): 111–140.

Biddle, G. C., and W. E. Ricks. 1988. Analyst forecast errors and stock price behavior near the earnings announcement dates of LIFO adopters. *Journal of Accounting Research* 26 (2): 169–194.

Boynton C., P. Dobbins, and G. Plesko. 1992. Earnings management and the corporate alternative minimum tax. *Journal of Accounting Research* 30 (Supplement): 131–154.

Chen, K., and M. Schoderbek. 2000. The 1993 tax rates increase and deferred tax adjustments: A test of functional fixation. *Journal of Accounting Research* 38 (1): 23–45.

———, M. Danielson, and M. Schoderbek. 2002. Analysts' interpretation of transitory earnings components: Evidence from forecast revisions after disclosure of the 1993 deferred tax adjustment. Working paper, Rutgers University.

Clement, M. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27 (3): 285–303.

Collins, J., and D. Shackelford. 1992. Foreign tax credit limitations and preferred stock issuances. *Journal of Accounting Research* 30 (Supplement): 103–124.

Conover, W. J. 1980. *Practical Nonparametric Statistics.* New York, NY: John Wiley & Sons.

Dhaliwal D., and S. Wang. 1992. The effect of book income adjustment in the 1986 alternative minimum tax on corporate financial reporting. *Journal of Accounting and Economics* 15 (1): 7–26.

Earley, P. C. 1985. Influence of information, choice and task complexity upon goal acceptance, performance, and personal goals. *Journal of Applied Psychology* 70 (3): 481–491.

Einhorn, H., R. Hogarth, and E. Klempner. 1977. Quality of group judgment. *Psychological Bulletin* 87 (1): 158–172.

Financial Accounting Standards Board (FASB). 2002. *Facts about FASB*. Stamford, CT: FASB.

Gramlich, J. 1991. The effect of the alternative minimum tax book income adjustment on accrual decisions. *The Journal of the American Taxation Association* 13 (1): 36–56.

Guenther, D. 1994. Earnings management in response to corporate tax rate changes: Evidence from the 1986 Tax Reform Act. *The Accounting Review* 69 (1): 230–243.

Hirst, E., and P. Hopkins. 1998. Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research* 36 (Supplement): 47–75.

———, ———, and J. Wahlen. 2002. Fair values, performance reporting, and bank analysts' risk and valuation judgments. Working paper, The University of Texas at Austin and Indiana University.

Iselin, E. 1988. The effects of information load and information diversity on decision quality in a structured decision task. *Accounting, Organizations and Society* 13 (2): 147–164.

Jacob, J. 1997. Discussion of do security analysts improve their performance with experience? *Journal of Accounting Research* 35 (Supplement): 159–166.

Kim, K., and D. Schroeder. 1990. Analysts' use of managerial bonus incentives in forecasting earnings. *Journal of Accounting and Economics* 13 (1): 3–23.

Maines, L., and L. McDaniel. 2000. Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial statement presentation format. *The Accounting Review* 75 (2): 179–207.

Manzon, G. 1992. Earnings management of firms subject to the alternative minimum tax. *The Journal of the American Taxation Association* 14 (2): 88–111.

———, and G. Plesko. 1999. Strategic disclosure for political gain: The case of the corporate alternative minimum tax. Working paper, Massachusetts Institute of Technology.

Maydew, E. 1997. Tax-induced earnings management by firms with net operating losses. *Journal of Accounting Research* 35 (1): 83–96.

McEwen, R., and J. Hunton. 1999. Is analyst forecast accuracy associated with accounting information use? *Accounting Horizons* 13 (1): 1–16.

Mikhail, M., B. Walther, and R. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* 35 (Supplement): 131–157.

———, ———, and ———. 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* 74 (2): 185–200.

O'Brien, P. C. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10 (1): 53–83.

Omer, T. and T. Atwood. 1991. Income reporting strategies of alternative minimum tax-based clienteles. Working paper, University of Illinois.

Paquette, L., and T. Kida. 1988. The effect of decision strategy and task complexity on decision performance. *Organizational Behavior and Human Decision Processes* 41 (1): 128–142.

Payne, J. 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance* 16 (2): 366–387.

———. 1982. Contingent decision behavior. *Psychological Bulletin* 92 (2): 382–402.

———, J. Bettman, and E. Johnson. 1988. Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14 (4): 534–552.

Philbrick, D., and W. Ricks. 1991. Using value line and I/B/E/S analyst forecasts in accounting research. *Journal of Accounting Research* 29 (2): 397–417.

Scholes, M., G. P. Wilson, and M. Wolfson. 1992. Firms' responses to anticipated reductions in tax rates: The Tax Reform Act of 1986. *Journal of Accounting Research* 30 (Supplement): 161–185.

Securities and Exchange Commission (SEC). 1997. *Plain English Disclosure*. 17 CFR Parts 228, 229, 230, 239 and 274. Release Nos. 33–7497, 34–39593, IC–23011, International Series No. 1113; File No. S7–3–97, RIN 3235–AG88. Washington, D.C.: Government Printing Office.

Smith, J., T. Mitchell, and L. Beach. 1982. A cost-benefit mechanism for selecting problem-solving strategies: Some extensions and empirical tests. *Organizational Behavior and Human Performance* 29 (3): 370–396.

Value Line. 1986a. *Value Line Investment Survey*. July 25. New York, NY: Value Line Publishing, Inc.

———. 1986b. *Value Line Investment Survey*. September 26. New York, NY: Value Line Publishing, Inc.

———. 1986c. *Value Line Investment Survey*. November 7. New York, NY: Value Line Publishing, Inc.

Wilkie, P. J., and S. Limberg. 1992. Measuring explicit tax (dis)advantage for corporate taxpayers: An alternative to average effective tax rates. *The Journal of the American Taxation Association* 14 (1): 76–91.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 297–325

The Influence of Tax and Nontax Factors on Banks' Choice of Organizational Form

Leslie Hodder
Stanford University
Mary Lea McAnally
Texas A&M University
Connie D. Weaver
The University of Texas at Austin

ABSTRACT: This paper identifies tax and nontax factors that influence commercial banks' conversion from taxable C-corporation to nontaxable S-corporation from 1997 to 1999, after a 1996 tax-law change allowed banks to convert to S-corporations for the first time. We find that banks are more likely to convert when conversion saves dividend taxes, avoids alternative minimum taxes, and minimizes state income taxes. Banks are less likely to convert when conversion restricts access to equity capital, nullifies corporate tax loss carryforwards, and creates potential penalty taxes on unrealized gains existing at the conversion date. Banks with significant deferred tax assets are less likely to convert, presumably because the write-off of deferred taxes at conversion decreases regulatory capital and exposes the bank to costly regulatory intervention. We also investigate the strategic choices banks make before converting to S-corporations. Converting banks alter their capital structures, deliberately sell appreciated assets, and strategically set dividends to augment net conversion benefits.

Keywords: *organizational form; S-corporation; regulatory costs; deferred tax.*

Data Availability: *Data are publicly available from sources identified in the paper.*

This paper benefited significantly from the comments of Shane Dikolli, Merle Erickson, Charlene Henderson, Eric Hirst, Ross Jennings, Bill Kinney, Lisa Koonce, Marc LeClere, Stephen Limberg, Tom Omer, John Robinson, Richard Sansing, Senyo Tse, Kristina Zvinakis, and seminar participants at the University of California, Davis, The Ohio State University, University of Oklahoma, Rice University, Texas A&M University, University of Toronto, University of Utah, and the 2001 American Accounting Association Annual Meeting. We are grateful for the comments and suggestions of two anonymous reviewers and Associate Editor Ken Klassen.

Submitted April 2001
Accepted August 2002

I. INTRODUCTION

This study examines tax and nontax factors that influence commercial banks' choice of organizational form, and explores the strategic choices banks make when they expect to change organizational form. Organizational form is important because it affects firms' taxes, financial reports, and valuation (Erickson and Wang 2000). Specifically, we examine banks' conversions from taxable C-corporations to nontaxable S-corporations after a 1996 tax-law change first permitted banks to become S-corporations. Because tax laws precluded banks from organizing as S-corporations prior to 1997, the 1996 tax-law change provides a natural experimental setting that allows us to directly assess factors that encourage firms to change organizational form. The discrete nature of the tax-law change allows us to document strategies that converting banks use to take advantage of a change in organizational form.

S-corporations differ from C-corporations only for tax purposes. That is, S-corporations retain the advantages of limited liability and unlimited life, but avoid the double tax burden imposed on C-corporations. C-corporations' earnings are taxed at the corporate level and then their dividends are also taxed at the individual-shareholder level. Under reasonable assumptions, a C-corporation that pays out all of its income as dividends has a maximum combined (corporate and individual) federal tax rate of up to 61 percent. In contrast, S-corporation earnings generally are not taxed at the corporate level and S-corporation dividends are tax-free. However, S-corporation shareholders are taxed individually on their *pro rata* shares of the company's earnings. The maximum combined federal (corporate and individual) tax rate for an S-corporation that pays out all of its income as dividends is 40 percent. Thus, the primary benefit of being an S-corporation is relief from double taxation. Additional benefits include tax relief both for banks that pay the alternative minimum tax and for banks in the 38 states whose tax laws recognize S-corporations.

Given the tax benefits of being an S-corporation, one might expect that many banks would convert. However, by the end of 1999, only 19.2 percent of banks had done so, presumably because converting to S-corporation imposes significant costs on banks. One potentially significant cost is that banks must write-off deferred tax assets when they convert to S-corporations (SFAS No. 109, para. 127, FASB 1992). Because bank regulatory capital includes retained earnings, eliminating deferred tax assets reduces banks' current-period regulatory capital. When regulatory capital slips below certain levels, banks face costly regulatory intervention. Furthermore, S-corporation rules restrict the number and type of shareholders, limiting S-corporations' ability to raise capital. Other costs arise because S-corporation income is taxed at higher individual rates, S-corporations cannot take advantage of net operating loss carryforwards, and they can incur penalty taxes on unrealized gains existing at conversion.

Our logistic regression results show that banks are more likely to convert to S-corporations as the tax benefits of conversion increase, and are less likely to convert as the costs of conversion increase. Specifically, banks are more likely to convert when conversion saves dividend taxes, avoids alternative minimum taxes, and minimizes state income taxes. At the same time, the likelihood of converting is negatively related to conversion costs, including restricted access to equity capital, loss of corporate tax-loss carryforwards, and potential penalty taxes on unrealized gains existing at the conversion date. We also find that banks with large deferred tax assets are less likely to convert, presumably because the required write-off of deferred tax assets decreases regulatory capital.

We find that banks employ two types of strategies before converting to S-corporations. First, many banks that were ineligible when Congress enacted the law in 1996 restructure their outstanding equity to become eligible to convert by repurchasing disallowed outstanding preferred shares and common shares held by foreign shareholders. At the extreme, some banks delist their publicly traded stock to conform to S-corporation restrictions. These surprising results suggest that there are few absolute *ex ante* disqualifications to converting to an S-corporation; banks convert when the tax savings from eliminating double taxation exceeds restructuring and other conversion costs.

As a second strategy, banks converting after 1997 (i.e., late converters) take actions to increase benefits or reduce costs before conversion. Specifically, net operating loss carryforwards sharply decline in the years prior to conversion, consistent with banks accelerating their use of net operating losses before conversion renders them unusable, or even delaying conversion until after they use net operating losses. In the years before conversion, banks selectively sell securities to recognize gains. This reduces potential penalty taxes on unrealized gains and accelerates banks' use of net operating loss carryforwards. Finally, in the years before becoming S-corporations, converting banks reduce dividend payments relative to nonconverting banks and relative to prior years. This strategy decreases the amount subject to double taxation in converting banks' last years as C-corporations. Taken together, these findings suggest that before converting to S-corporations, banks make operating and financing changes to minimize the costs and maximize the benefits of the conversion.

Our study contributes to the literature in a number of ways. First, prior research provides mixed evidence as to whether taxes *significantly* influence corporate decisions (Amihud et al. 1990; Ayers et al. 1996; Klassen 1997). In contrast, we report that specific tax benefits affect organizational-form choices in spite of the significant costs associated with conversion. Second, Congress is currently considering proposals to liberalize restrictive provisions on the number and type of shareholders for S-corporations, including banks. Our results identify additional tax and nontax factors that future legislation could address. For example, future legislation might consider eliminating penalty taxes on unrealized gains as a way to encourage conversions. Third, we identify explicit strategies firms use to capture potential tax savings. To our knowledge, extant research has not previously documented this sort of anticipatory strategic behavior in a tax setting. Fourth, we document a significant and previously unidentified regulatory consequence of deferred tax accounting—financial reporting of deferred taxes is associated with operating decisions such as organizational form choice. Finally, a recent U.S. General Accounting Office (GAO) report discusses bank regulators' concerns about how proposals to liberalize S-corporation rules could affect the community banking system (U.S. GAO 2000). By identifying the specific factors that appear to affect banks' conversion decisions, our study could help bank regulators assess how S-corporations affect community bank safety, soundness, and liquidity.

II. PRIOR RESEARCH

Prior studies yield mixed evidence on whether taxes affect firms' organizational form choices. Conclusions depend on the specific organizational form choice (sole proprietorship, partnership, S-corporation, or C-corporation) and the research design (use of public or private firms and data-measurement timing).

Studies that examine the organizational-form choices of publicly traded partnerships and corporations find, in most cases, that firms choose organizational forms to reduce taxes (Guenther 1992; Terando and Omer 1993; Shaw and Wier 1993; Gentry 1994; Damodaran et al. 1997; Omer and Terando 1999). However, other financial *nontax* factors, including limited liability, debt levels, cash flow, and dividend and nondividend distributions, can also affect organizational form choice. In particular, researchers have struggled to discern the benefits of limited liability. In contrast, our examination of banks' conversions from C-corporations to S-corporations holds constant the effect of limited liability because regulatory agencies require banks to be corporations.

A second stream of studies considers *private* entities' organizational form choices. Here, results concerning the effect of taxes are mixed, perhaps because researchers cannot observe an entity's attributes until after it makes its organizational form choice. Ayers et al. (1996) examine survey data for a sample of existing C-corporations, S-corporations, and unincorporated entities and find that nontax factors such as business risk, ownership structure, firm size, and age are associated with organizational form choices, but that taxes are not. In contrast, using proprietary Internal Revenue Service (IRS) data, Plesko (1994) finds that taxes are associated with organizational form choice for newly incorporated companies. However, measuring firm attributes after the firm chooses its organizational form does not distinguish whether the observed attributes determined the entity's organizational form or vice versa. As Plesko (1994, 317) states, "To analyze the decision of choice would require access to detailed information about the firms at the time the choice decision was being considered."

Plesko (1995) and Omer et al. (2000) use confidential tax return data measured before the organizational form choice to evaluate conversions of firms that were eligible to become S-corporations before the Tax Reform Act of 1986 (TRA86, U.S. Congress 1986), but chose not to. Thus, they examine whether the TRA86 tax-law changes altered the relative costs and benefits of S-corporations. Plesko (1995) reports a surge in conversions after TRA86, presumably because the decrease in personal tax rates relative to corporate rates made the S-corporation form more attractive. However, fewer firms than expected converted to S-corporations, and conversion rates varied dramatically by industry. For example, Omer et al. (2000) find that potential tax penalties on unrealized gains after TRA86 discourage S-corporations conversions in the natural resource industry, where these gains are potentially high.

In contrast to Plesko (1995) and Omer et al. (2000), who study firms that were eligible to become S-corporations before TRA86 but chose not to, we study firms facing a new organizational-form opportunity set created when tax-law changes removed two restrictions that precluded banks from organizing as S-corporations prior to 1997. First, on August 20, 1996, the Small Business Job Protection Act (P.L. 104-188, U.S. Congress 1996) removed the restriction disallowing banks that use the reserve method for loan losses from organizing as S-corporations.[1] Second, the IRS automatically revokes the S-status of corporations whose passive investment income exceeds 25 percent of gross receipts for three consecutive years (IRC §1362(d)(3)).[2] An IRS ruling, issued December 20, 1996, removed this barrier by redefining most types of interest income banks earn as "active income" (Notice 97-5,

---

[1] Prior tax law stated that financial institutions were ineligible if (1) they used the reserve method of accounting for loan losses or (2) they could have used the reserve method had the size limitation in § 585 not prohibited them from doing so. Internal Revenue Code (IRC) § 1361(b)(2) prohibits banks with assets greater than $500 million from using the reserve method. Together, these requirements effectively precluded all financial institutions from electing the S-corporation form.

[2] This provision applies to S-corporations that have subchapter C earnings and profits. Because banks were taxed as C-corporations prior to the law change, this provision applied to all converting banks. This was particularly restrictive because a large proportion of banks' income consists of interest and dividends, which were considered passive-investment income prior to the change (Notice 97-5, 1997-1 CB 352, IRS 1996).

1997-1 CB 352, IRS 1996). Thus, banks first became eligible to chose the S-corporate form after 1996.[3]

III. DATA, MODEL, AND VARIABLE DEFINITIONS

Data

Our sample consists of 6,622 private commercial banks that filed a Report of Condition and Income (call report) with the Federal Reserve or Federal Deposit Insurance Corporation (FDIC) during the period 1993–1999.[4] Of these, 1,272 banks converted to S-corporations by the end of 1999 and 5,350 continued to operate as taxable C-corporations. We identify 575 conversions to S-corporate form (9 percent of the total sample) by the end of 1997, 1,013 conversions (15 percent) by December 1998, and 1,272 conversions (19 percent) by the end of 1999.

Call reports provide data at the bank operating level, which allows us to include all private banks in our sample and ensures consistent data across banks regardless of whether a holding company owns the bank.[5] S-corporation rules require that banks have 75 or fewer shareholders. Because call report data do not reveal the number of shareholders for each bank, we are unable to verify whether a bank exceeds the 75-shareholder maximum. However, previous research indicates that most private companies have far fewer than 75 shareholders (Plesko 1994; Ayers et al. 1996). Furthermore, private banks with more than 75 shareholders may take action to reduce the number of shareholders to conform to the S-corporation limit, in anticipation of converting. Consequently, we do not restrict the sample to banks with 75 or fewer shareholders; however, we do consider all publicly traded banks as ineligible to convert to S-corporation and exclude them from the sample.[6]

Model

We define our dependent variable (S) as 1 if a bank converted to an S-corporation during the sample period (1997–1999), and 0 otherwise. Our model includes six tax variables (DIVIDEND, TAXINC, BUILTIN, LOSSCF, AMT, and STATE) and three nontax variables (GROWTH, REGULATORY, and LEVERAGE). Consistent with prior research, we also include bank size and age as control variables.

$$\begin{aligned} S_t = \alpha &+ \beta_1 \mathrm{DIVIDEND}_{94\text{–}96} + \beta_2 \mathrm{TAXINC}_{94\text{–}96} + \beta_3 \mathrm{BUILTIN}_{t-1} + \beta_4 \mathrm{LOSSCF}_{t-1} \\ &+ \beta_5 \mathrm{AMT}_{94\text{–}96} + \beta_6 \mathrm{STATE}_{t-1} + \beta_7 \mathrm{GROWTH}_{94\text{–}96} + \beta_8 \mathrm{REGULATORY}_{t-1} \\ &+ \beta_9 \mathrm{LEVERAGE}_{94\text{–}96} + \beta_{10} \mathrm{SIZE}_{94\text{–}96} + \beta_{11} \mathrm{AGE}_{99}, \end{aligned} \tag{1}$$

---

[3] To become S-corporations, firms must gain unanimous shareholder consent. Under the new law, to convert to S-corporations in 1997, banks had to file the S-election forms by March 17, 1997. Banks converting in 1997 had approximately three months (December 20, 1996–March 17, 1997) to meet the eligibility requirements and obtain shareholder consent.

[4] To enter the sample, banks must file Call Reports from 1993 through 1999 so that we may calculate the lead and lag variables in our empirical models.

[5] S-corporation rules allow a bank holding company to make an S-election for itself along with qualified subchapter-S subsidiary elections for its subsidiary banks. A subsidiary bank may itself become an S-corporation provided that (1) the subsidiary is 100 percent owned by the bank holding company and (2) the bank holding company also becomes an S-corporation. Tax rules treat the qualified subchapter S subsidiary bank's assets, liabilities, income, and deductions as those of the holding company.

[6] Excluding publicly traded banks reduces, but does not eliminate, the possibility that our sample includes banks with more than the maximum 75 shareholders. This works against our hypotheses, because even though the sample banks' data may suggest they should convert, they do not because the (unobserved) number of shareholders exceeds 75.

where:

DIVIDEND = dividends scaled by assets, 1994–1996 average;
TAXINC = taxable income scaled by assets, 1994–1996 average;
BUILTIN = market value less book value of investment securities available for sale at the end of the year prior to conversion, scaled by assets;
LOSSCF = 1 if bank has tax loss carryforwards at the end of the year prior to conversion, 0 otherwise;
AMT = a proxy for the bank's alternative minimum tax, defined as tax-exempt income scaled by assets, 1994–1996 average;
STATE = 1 if bank operates in a state where state taxes render conversion most beneficial, 0 otherwise;
GROWTH = annual percentage change in assets, 1994–1996 average;
REGULATORY = 1 if a deferred tax write-off would inflict significant adverse regulatory consequences, 0 otherwise;
LEVERAGE = nondeposit liabilities scaled by assets, 1994–1996 average;
SIZE = natural log of assets (in thousands), 1994–1996 average; and
AGE = natural log of number of years since incorporation, as of 1999.

In the window between the tax-law change and the bank's conversion date, bank management can manipulate firm attributes to maximize *ongoing* benefits (or minimize ongoing costs) of conversion. For example, by reducing (or even eliminating) dividends in the year immediately before conversion, a high-dividend-paying bank can increase the tax benefits of converting to an S-corporation.[7] Using the temporarily lowered dividends as an independent variable would obscure the underlying attribute (high-dividend-paying bank). Because our research question pertains to the underlying attributes associated with banks' decisions to convert, we are interested in unmanaged values of those attributes. Consequently, we measure ongoing costs and benefits prior to the earliest possible S-corporation conversion in 1997 to avoid contaminating our tests with discretionary changes designed to increase the net benefits of converting. Specifically, we measure DIVIDEND, TAXINC, AMT, GROWTH, LEVERAGE, and SIZE using the average over the three years preceding the tax-law change, and denote this with the subscript "94–96" in Equation (1).

Other variables that capture *one-time* costs are more meaningfully measured at the conversion date. Specifically, the tax code defines built-in gains as those existing on the date of conversion. Similarly, a bank loses the benefit of net operating loss carryforwards that exist at conversion, and potential regulatory costs depend on the deferred tax assets at conversion. Thus, we measure BUILTIN, LOSSCF, and REGULATORY immediately before conversion, consistent with the one-time nature of these costs. For the nonconverters, we measure one-time costs at the end of 1998.[8] We measure STATE in the year preceding the conversion year because banks' location does not vary substantially over our time period. We measure AGE in 1999. We explain each of our independent variables below.

---

[7] Distributions to S-corporation shareholders are tax-free to the extent that the distributions represent income the firm earned while operating as an S-corporation. If distributions exceed cumulative S-corporation income and if the firm was previously a C-corporation, the excess distributions are generally taxed as dividends from a C-corporation. In our sample, banks that convert to S-corporations pay about 40 percent of annual income in dividends during the 1994–1996 period. Because these average payouts are substantially less than 100 percent, the distribution rules do not severely restrict a bank's ability to reduce dividend payments before conversion and then increase dividends afterward.

[8] To ensure that using 1998 data for lagged variables for nonconverters does not drive our results, we also run the logistic analysis using 1997 data for the nonconverters. Our inferences are unchanged.

Tax-Related Determinants of the S-Corporation Conversion Choice

Dividend Taxes and Taxable Corporate Income

The numerical analysis in Figure 1 shows that the tax benefit of S-corporations (excluding effects of built-in gains, tax loss carryforwards, and alternative minimum tax) has two components:

$$\text{Tax Benefit} = \text{Dividends} \times \tau_i + \text{Taxable income} \times (\tau_c - \tau_i). \qquad (2)$$

The first part of the tax benefit arises from eliminating individual-level taxes on dividends, and increases with the magnitude of dividends. The second part of the benefit arises from the difference between the corporate tax rate (τ_c) and the individual tax rate (τ_i). These two

FIGURE 1
Numerical Illustration of S-Corporation Tax Benefits

	Dividend-Paying Bank[a]	*Non-Dividend-Paying Bank*[b]
Before Bank Converts to an S-Corporation		
Corporate taxable income (TI)	$1,000	$1,000
Tax on corporate taxable income ($\tau_c \times$ TI)	(350)	(350)
Corporate income after-tax	650	650
Dividend paid to individual shareholder	650	0
Tax on dividend ($\tau_i \times$ D)	(260)	0
Individual income after-tax	390	650
Total taxes paid	$610	$350
After Bank Converts to an S-Corporation		
Corporate taxable income (TI)	$1,000	$1,000
Tax on corporate taxable income ($\tau_c \times$ TI)	0	0
Corporate income after-tax	1,000	1,000
Dividend paid to individual shareholder	1,000	0
Tax on dividend ($\tau_i \times$ D)	(400)	(400)
Individual income after-tax	600	600
Total taxes paid	$400	$400
Benefit of bank converting to an S-corporation	$210	($50)

Assumptions underlying the numerical illustration:
Corporate tax rate $\tau_c = 35\%$;
Individual tax rate $\tau_i = 40\%$;
Taxable income TI = $1,000; and
Dividend D = 100% of after-tax corporate income.

[a] Before converting to an S-corporation, the bank pays total taxes of $610 or 61 percent of taxable income. If the bank converts, then it saves taxes of $210 and increases after-tax income by the same amount. The benefit of conversion increases with the magnitude of dividends paid.

[b] Before converting to an S-corporation, the non-dividend-paying bank pays total taxes of $350 or 35 percent of taxable income. If the bank converts, then taxes are $400, $50 higher. The assumption that individual tax rates exceed corporate tax rates ($\tau_i > \tau_c$), as was the case during the 1997–1999 time period covered by our study, drives this result. The benefit of an conversion decreases with the magnitude of taxable income.

tax rates do not change over the 1996–1999 time period covered by our study, and maximum individual tax rates exceed corporate tax rates ($\tau_c < \tau_i$). Thus, tax benefits to S-corporations decrease with banks' taxable income.

Because the two components (dividends and taxable income) work in opposite directions, we distinguish them to test separately their influence on banks' organizational form choice. Prior research did not separate the two components, presumably because during the time periods studied, corporate tax rates exceeded individual tax rates ($\tau_c > \tau_i$), making the dividend and taxable income components additive (Plesko 1995; Ayers et al. 1996; Omer et al. 2000). In supplementary analysis, we include a combined tax benefit variable in our model.

We measure the first tax component, DIVIDEND, as average dividends paid (as a percentage of assets) over the 1994–1996 period.[9] We predict that the probability that banks convert to S-corporations increases with DIVIDEND because the more dividends banks pay, the greater the benefit from converting. We include a measure of taxable income, TAXINC, to capture the second component. To calculate TAXINC, we use the current tax provision from the income statement, which we gross up by the highest marginal corporate tax rate of 35 percent. We scale TAXINC by assets and average over the 1994–1996 period.[10] Because the tax rate differential is negative ($\tau_c - \tau_i < 0$), we expect the coefficient on TAXINC to be negative. That is, we expect that the probability that banks convert decreases with banks' taxable income.

Built-In Gains (Tax Penalty on Unrealized Gains Existing at Conversion)

Gains that exist at the time a bank becomes an S-corporation (i.e., built-in gains) are taxed at the corporate level as well as the individual shareholder level if the gains are *realized* within ten years. The high turnover in banks' securities portfolios makes it likely that the bank will realize built-in gains within the ten-year window. The rationale for taxing an S-corporation bank on realized built-in gains is that the bank "earned" the gains while it was a taxable C-corporation. Thus, S-corporations and their shareholders must immediately pay double income taxes (corporate and individual) when built-in gains are realized. In contrast, C-corporation shareholders do not pay individual tax on realized gains until the bank pays them out as dividends. Consequently, C-corporation banks can indefinitely defer individual taxes on realized gains by withholding dividends. In summary, conversion curtails a bank's ability to defer individual-level taxes on realized built-in gains, which in turn reduces the benefits of S-corporation conversion.

The size and price volatility of banks' marketable securities portfolios create the potential for substantial built-in gains. Thus, we define BUILTIN as the excess of the market value over the book value of banks' available-for-sale securities portfolio at the end of the year prior to conversion (scaled by total assets) to capture this one-time cost of conversion.

---

[9] Omer et al. (2000) do not include a separate measure of the dividend tax savings in their model. In sensitivity tests, they measure the dividend tax savings as the product of the highest individual tax rate and the proportion of income the firm paid out as dividends. Our dividend tax savings variable differs in that we eliminate the scalar multiplier (tax rates) and scale bank dividends by assets rather than income. In sensitivity tests, we scale dividends by pretax income and find similar results to those reported here, although we lose 196 banks from the sample because they reported zero or negative income.

[10] Omer et al. (2000) measure the tax savings as a corporation's gross tax liability minus credits (from the tax return), grossed up by the top corporate marginal tax rate and multiplied by the individual/corporate tax rate difference. Although we measure TAXINC in the spirit of Omer et al. (2000), we do not have access to tax return data. Therefore we proxy the bank's gross tax liability with the current tax provision from the income statement. Additionally, we scale TAXINC by assets to measure the size of the tax cost relative to the total size of the bank.

Because only built-in gains trigger corporate-level tax for an S-corporation, we code BUILTIN as 0 when the available-for-sale portfolio has unrealized losses. We exclude unrealized gains on held-to-maturity securities because built-in-gains taxes apply only if the bank sells appreciated securities within ten years—and, by definition, management does not intend to sell held-to-maturity securities. In sum, BUILTIN reflects *potential* double taxes because banks will not necessarily sell securities available for sale within ten years. We expect the coefficient on BUILTIN to be negative because the greater the potential built-in gains taxes, the less likely the bank is to convert to an S-corporation.

Tax Losses

C-corporations can carry tax losses back two years and forward 20 years to offset *corporate* taxable income. However, S-corporations lose the carryforward benefit. Additionally, the years the firm operates as an S-corporation deplete the 20-year-carryforward period. Thus, banks with a net operating loss carryforward have less incentive to convert to S-corporations.[11]

Call reports do not provide net operating loss carryforward information. Therefore, we define an indicator variable (LOSSCF) equal to 1 if the bank's prior year's income tax provision either (1) is 0 or (2) is negative and the negative amount cannot be fully absorbed by the previous two years' income tax provisions (i.e., the sum of the current negative provision and the sum of the previous two years' provisions is negative); otherwise, we set LOSSCF equal to 0. LOSSCF therefore represents the existence (but not the magnitude) of net operating loss carryforwards that the bank would lose upon corporation conversion, so we predict that banks with loss carryforwards are less likely to convert.

Alternative Minimum Tax

Banks frequently pay alternative minimum tax (AMT) because they hold extensive portfolios of tax-exempt securities that generate income subject to AMT. S-corporation conversion confers a tax benefit to banks that pay AMT because S-corporations are not subject to the corporate AMT.[12] To avoid paying AMT, banks could simply not invest in tax-exempt securities. However, banks hold tax-exempt assets for nontax reasons. For example, banks pledge tax-exempt securities as collateral against governmental deposits or hold securities of local municipalities to strengthen community relations. Therefore, C-corporation banks cannot uniformly avoid investing in tax-exempt securities to avoid paying AMT. In fact, the GAO reports that from 1987 to 1992, financial institutions paid $1.2 billion of AMT per year—about 25 percent of the total corporate AMT collected in those years (U.S. GAO 1995). We define AMT as the bank's 1994–1996 average income from tax-exempt securities, loans, and leases, less tax-exempt interest expense, scaled by total assets.[13] We predict that the more AMT a bank pays, the more likely the bank is to convert to an S-corporation.

---

[11] Future operating losses have an opposite effect on the conversion decision. Banks that expect to suffer continued losses for an extended period (Ayers et al. 1996) could view S-status favorably because S-corporation losses flow through to shareholders who can use the losses to offset ordinary income. When individual tax rates exceed corporate rates, S-corporation status would be attractive for banks that expected to incur continued losses. However, very few banks experienced continued operating losses during our sample period.

[12] Individuals separately determine AMT. However, individual AMT rules do not include general tax-exempt interest (the item that frequently causes banks to pay AMT) as a preference item or adjustment for the AMT calculation.

[13] Omer et al. (2000) measure AMT as one-half the difference between book and taxable income. We lack the necessary confidential tax return data to replicate the Omer et al. (2000) measure.

State Tax Considerations

Prior research demonstrates that state taxes affect firms' tax strategies (Petroni and Shackelford 1995; Sawyers and Beasley 1998; Beatty and Harris 2001a, 2001b). We identify two issues that affect a bank's potential state tax savings from converting to an S-corporation. First, ten states do not recognize S-corporations. These states impose a state-level corporate tax regardless of organizational form. Banks that do *not* operate in one of these ten states are more likely to convert to S-corporations to avoid double taxation at the state level (Plesko 1994). Second, states with unitary rules identify bank holding companies and their subsidiary banks as one entity for state reporting purposes and require them to file combined reports. Combined reporting is beneficial because the bank holding company can offset the losses of unprofitable subsidiary banks against the earnings of profitable subsidiary banks. In states that do not allow combined reporting, each subsidiary bank must report its full taxable income, rather than allocating combined income among the states. Because an S-corporation mimics combined reporting, banks that operate in states that do not allow combined reporting receive an additional benefit if they are S-corporations.

We define a dichotomous variable STATE to represent the most favorable state tax provisions for an S-corporation conversion: banks with STATE = 1 operate in states that recognize S-corporations and *do not* allow combined reporting.[14] We expect the coefficient on STATE to be positive.

Nontax-Related Determinants of the S-Corporation Conversion Choice

Demand for External Capital

A disadvantage of converting to an S-corporation is that it limits the supply of external equity. Tax laws restrict S-corporations to a maximum of 75 shareholders, and prevent partnerships, corporations, and other limited liability entities from owning S-corporation shares. Consequently, S-corporation banks cannot attract institutional investors or sell shares publicly. In addition, tax laws mandate that S-corporations issue only one class of stock. The cost of these disadvantages increases with a bank's demand for equity. We use historical growth as a proxy for the bank's demand for equity. The higher a bank's growth rate, the more equity it will need to take advantage of profitable lending or deposit-taking opportunities while maintaining regulatory capital ratios. We define GROWTH as the average percentage change in the bank's total assets over the 1994–1996 period, and we expect that conversion to S-corporation is less likely for faster-growing banks that demand more capital.

Regulatory Cost

SFAS No. 109 (para. 127) requires that a firm write off deferred taxes to income when the firm's tax status changes. Because eliminating deferred tax *assets* from the balance sheet reduces regulatory capital, banks with significant deferred tax assets relative to regulatory capital have less incentive to become S-corporations. S-corporation conversion also has long-term regulatory consequences; following conversion, regulatory capital increases annually by pretax earnings because S-corporations have no tax expense. However, S-corporation banks will likely increase dividends after converting, to cover the higher taxes shareholders must pay directly on their shares of the S-corporation profits.[15] Moreover, S-corporation banks cannot accrue deferred tax benefits arising from future deductions against corporate income because the S-corporation bank will not receive such future benefits. Thus, the accounting for deferred taxes has adverse immediate and long-term consequences for S-corporations' regulatory capital.

---

[14] The intersection of these two conditions results in STATE = 1 for: Alabama, Arkansas, Delaware, Florida, Iowa, Maryland, Massachusetts, Mississippi, New Jersey, Oklahoma, Rhode Island, South Carolina, Vermont, and Wisconsin.

Regulatory capital is important because regulators use a simple but rigid decision rule: when regulatory capital falls toward or below statutory minimums, the bank is subject to increasingly stringent (and costly) regulatory intervention. In addition, the FDIC charges risk-related deposit insurance premiums based on banks' capital levels.[16] Moving to a lower capital level increases deposit insurance premiums even if the bank is above the minimum regulatory capital level.

We use the following three criteria to identify banks for which a deferred tax asset write-off could inflict significant adverse present and future regulatory consequences: (1) banks that have significant deferred tax assets relative to regulatory capital (in the 95th percentile of deferred tax assets to regulatory capital in the year prior to conversion); (2) banks that drop one or more regulatory capital levels when they write off deferred taxes; and (3) banks that fall below the minimum 8 percent risk-based capital level when they write off deferred taxes. We define a categorical variable REGULATORY as 1 if a bank meets any of the three criteria.

REGULATORY excludes deferred tax assets (liabilities) arising from unrealized losses (gains) on available-for-sale securities, to be consistent with regulatory capital rules. We also exclude from REGULATORY net deferred tax liabilities because banks converting to S-corporations do not necessarily write off deferred tax liabilities. We predict a negative coefficient on REGULATORY because a bank is less likely to convert to an S-corporation if conversion would stimulate adverse regulatory consequences.

The regulatory effect of the deferred tax asset write-off is potentially confounded by AMT payments that create tax credits that a bank can use in future periods to offset its regular tax liability when it exceeds its AMT liability. Banks account for these tax credits as deferred tax assets and converting banks will write off such deferred tax assets. Our data do not permit us to evaluate the existence or magnitude of these credits. Although our tests do not explicitly consider forgone AMT credits, we believe they represent a modest deterrent to S-conversion because, in contrast to net operating loss carryforwards, AMT tax credits do not expire.

Leverage

We include a measure of leverage to capture the effects of capital structure on banks' conversion choice. We measure LEVERAGE as nondeposit liabilities divided by total assets. To the extent that more highly levered banks use tax-deductible interest to reduce their corporate tax bills, they are less likely to need the additional costly tax-planning strategy conversion offers. Cloyd et al. (1997) and Omer et al. (2000) document a negative relation between leverage and S-corporations. However, Plesko (1994) finds that S-corporations are associated with higher leverage, presumably because S-corporations have greater flexibility in structuring shareholder loans relative to C-corporations. This flexibility increases the benefit of conversion. Given this conflicting prior evidence, we make no prediction on the sign for the LEVERAGE coefficient.

---

[15] S-corporations need only distribute an amount equal to the individual tax rate times the difference between corporate taxable income and distributed dividends to make shareholders indifferent to a S-corporation conversion. Future research could investigate how S-corporations and their shareholders respond to the net tax savings that this organizational form provides.

[16] The Federal Reserve Bank considers a bank well capitalized if the risk-adjusted capital ratio is greater than 10 percent, adequately capitalized if the ratio is between 8 percent and 10 percent, undercapitalized if the ratio is between 6 percent and 8 percent, significantly undercapitalized if the ratio is between 2 percent and 6 percent, and critically undercapitalized if the ratio is less than 2 percent.

Control Variables for the S-Corporation Conversion Choice

Size

We include bank size to proxy for transaction costs or potential tax benefits. We measure SIZE as the natural log of total assets. Ayers et al. (1996) predict that large firms find it more costly to convert because of their operational complexity and multiple levels of management; however, their empirical tests do not reveal a significant relation between firm size and conversions. In contrast, Omer et al. (2000) predict and find that conversion benefits increase with size, presumably because of graduated corporate tax rates. In light of these conflicting results, we do not predict the direction of the relation between bank size and conversions.

Firm Age

Older banks may have established organizational efficiencies or entrenched management, both of which decrease the likelihood of conversion (Omer et al. 2000). Conversely, older banks may have shareholders who seek larger distributions and who would prefer a tax-saving conversion. Moreover, S-corporation conversion could restrict younger banks' growth opportunities (Ayers et al. 1996). Given this conflicting prior evidence, we do not predict the sign of the AGE coefficient. A number of banks in the sample are very mature. For example, the bank at the 90th percentile is 92 years old and the most established bank is 193 years old. To mitigate the effects of such extreme values, we define AGE as the natural log of the number of years since the bank was incorporated, as of 1999.

IV. RESULTS OF MODEL ESTIMATION

Descriptive Univariate Statistics

Table 1 reports descriptive statistics for banks that convert to S-corporations and banks that do not convert by 1999. Panel A includes statistics for firm attributes, and Panel B describes our regression variables. Panel A reveals that the average S-corporation is smaller than the average C-corporation (assets of $61 vs. $108 million), and reports, on average, less pretax income ($1.2 million compared to $1.9 million). In the years before conversion, banks that convert pay out on average 41.3 percent of income as dividends, compared to only 30 percent for banks that remain C-corporations.

Panel B indicates that the independent variables differ significantly across the subsamples of banks that convert and banks that do not convert, in directions consistent with our predictions. On average, converting banks have higher dividend payout ratios, lower taxable income, smaller built-in gains subject to the penalty tax, and are less likely to have loss carryforwards that would be lost upon conversion. Converting banks' tax savings from avoiding the AMT are greater and they are more likely to operate in a state where conversion is most beneficial. As expected, converting banks are slower growing and are less likely to suffer adverse regulatory consequences upon conversion.

Table 1, Panel C reports additional descriptive statistics for deferred tax measures. On average, deferred tax assets are a significantly smaller proportion of regulatory capital for converting banks (0.01 percent) than for nonconverting banks (1.44 percent). The observations in the upper tail of the distribution of deferred tax assets to regulatory capital include a disproportionate number of nonconverting banks. For example, the lower part of Panel shows that 5.5 percent of nonconverting banks fall above the 95th percentile of the distr bution of deferred tax assets to regulatory capital, whereas less than 3 percent of converti

TABLE 1

Descriptive Statistics for Firm Attributes, Logistic Regression Variables, and Deferred Tax Variables for Banks That Convert to S-Corporations Compared to Banks That Do Not Convert

Panel A: Firm Attributes

	Banks that Convert in 1997–1999 n = 1,272			*Banks that Do not Conver by 1999 n = 5,350*		
Variable	*Mean*	*Med.*	*Std. Dev.*	*Mean*	*Med.*	*Std. De*
Tax						
Dividends (in $000)	446.23	241.67	241.67	584.51	144.33	5,234.3
Dividends/income	41.3%	35.3%	44.9%	30.0%	22.6%	58.1
Current tax expense (in $000)	396.54	222.17	1,440.48	576.48	194.67	4,784.2
Built-in gains (in $000)	49.45	10.33	270.26	66.54	7.33	456.6
Tax-exempt income (in $000)	167.76	85.50	260.53	198.81	69.00	782.9
Nontax						
Deferred tax asset (in $000)	95.57	1.93	423.32	456.18	31.38	5,473.5
Deferred tax liability (in $000)	35.62	0.00	84.01	129.77	0.00	3,386.0
Non-deposit liabilities (in $000)	0.03	0.001	0.40	0.04	0.00	0.9
Average assets (in $000)	61,646	43,020	77,629	107,895	47,868	526,70
Pre-tax income (in $000)	1,194	756	1,718	1,940	719	11,97

Panel B: Independent Logistic Regression Variables

	Banks that Convert in 1997–1999 n = 1,272			*Banks that Do not Convert by 1999 n = 5,350*		
Variable	*Mean*	*Med.*	*Std. Dev.*	*Mean*	*Med.*	*Std. De*
Tax						
DIVIDEND	0.009*	0.008*	0.006	0.006	0.004	0.021
TAXINC	0.012*	0.011*	0.008	0.016	0.014	0.073
BUILTIN	<0.001*	<0.001*	0.001	0.001	0.001	0.003
LOSSCF	0.030*	0.000*	0.170	0.068	0.000	0.252
AMT	0.003*	0.002*	0.002	0.002	0.002	0.002
STATE	0.215*	0.000*	0.411	0.189	0.000	0.392
Nontax						
GROWTH	0.134*	0.007*	0.492	0.344	0.044	2.771
REGULATORY	0.033*	0.000*	0.179	0.062	0.000	0.242
LEVERAGE	0.029*	0.014*	0.034	0.031	0.013	0.066
Control						
SIZE	10.764*	10.737*	0.781	10.892	10.843	0.989
AGE	4.102*	4.407*	0.732	3.906	4.330	0.925

(*continued on next pag*

TABLE 1 (continued)

Panel C: Statistics for Additional Deferred Tax Variables

	Banks that Convert in 1997–1999 n = 1,272			*Banks that Do not Convert by 1999 n = 5,350*		
Variable	*Mean*	*90th Percentile*	*95th Percentile*	*Mean*	*90th Percentile*	*95th Percentile*
Deferred tax assets (in $000)	96	222	440	456	478	878
Deferred tax liabilities (in $000)	36	110	187	130	111	204
Regulatory capital (in $000)	7,602	14,746	20,635	16,732	22,654	35,463
Risk-based capital ratio	18.74%	27.35%	33.04%	19.62%	29.07%	35.59%
Deferred tax/regulatory capital	0.01%	3.00%	4.25%	1.44%	3.89%	5.44%
If deferred tax assets written-off:						
Number (%) of banks below 8% minimum risk-based capital ratio	3 (0.24)			21 (0.39)		
Number (%) of banks that fall one or more capital ratio categories	7 (0.55)			76 (1.42)		
Total banks below 8% or fall one or more categories	10 (0.79)			97 (1.81)		
Number (%) of banks with deferred tax assets to regulatory capital in 95th percentile	38 (2.99)			294 (5.50)		
Number (%) of banks with deferred tax assets to regulatory capital in 90th percentile	96 (7.55)			595 (11.12)		

---

* Differences between converting and nonconverting banks are significant at a 0.10 level or less, based on two-sample t-tests and Wilcoxon signed-rank tests.

Variable Definitions:

DIVIDEND = dividends scaled by assets, 1994–1996 average;

TAXINC = current tax provision divided by the highest corporate tax rate (35 percent), then scaled by assets, 1994–1996 average;

BUILTIN = market value less book value of investment securities available for sale, scaled by average assets, measured in the year prior to conversion (1998 for nonconverting banks);

LOSSCF = 1 if the prior year's income tax provision is zero or the prior year's income tax provision is negative and cannot be fully absorbed by the previous two years' income tax, and 0 otherwise. LOSSCF is measured in 1998 for nonconverting banks;

AMT = a proxy for the bank's alternative minimum tax, defined as income from tax-exempt leases and securities, scaled by assets, 1994–1996 average;

STATE = 1 if bank has a holding company *and* is located in a state without combined reporting *and* S-corporations are recognized for state purposes, and 0 otherwise, measured in the year prior to conversion (1998 for nonconverting banks);

GROWTH = annual percentage change in total assets, 1994–1996 average;

(*continued on next page*)

TABLE 1 (continued)

REGULATORY = 1 if bank is above 95th percentile of deferred tax assets to regulatory capital, *or* if writing off deferred tax assets prior to conversion causes bank to fall one or more regulatory capital categories, *or* to fall below the 8 percent regulatory capital minimum, and 0 otherwise, measured in the year prior to conversion (1998 for nonconverting banks);
LEVERAGE = nondeposit liabilities to assets, 1994–1996 average;
SIZE = natural log of assets (in thousands), 1994–1996 average;
AGE = natural log of number of years since incorporation as of 1999;
Regulatory capital = risk-weighted Tier I and Tier II capital as defined by Federal Reserve Bank regulation, measured in the year prior to conversion (1998 for nonconverting banks); and
Risk-based capital ratio = regulatory capital divided by total risk-adjusted asset as defined by Federal Reserve Bank regulation, measured in the year prior to conversion (1998 for nonconverting banks).

banks fall in that range. The consequences of writing off deferred tax assets differ across the converter and nonconverter subsamples: 107 banks would either fall one regulatory capital level (see footnote 16) or be undercapitalized if they wrote off their deferred tax assets. Only ten of these banks converted to S-corporations; 97 remained C-corporations.

Logistic Regression Results

We report logistic regression results for the likelihood a bank will convert to an S-corporation in Table 2. Our pooled regression includes 1,272 banks that converted at any time between 1997 and 1999 and 5,350 banks that did not convert by the end of 1999. Overall, the model correctly predicts 70 percent of the conversions.[17] We report the elasticity of each variable to calibrate its effect on banks' overall probability of converting to S-corporations. Elasticities represent the responsiveness of the probability of conversion to small (we use 1 percent) changes in the independent variables. In Table 2, we also report sensitivity measures of the change in the probability of conversion when continuous variables change from the 25th percentile value to the 75th percentile, and each categorical variable changes from 0 to 1 (LeClere 1992).

Tax Variables

Table 2 shows that our tax factors all have the expected relation to banks' likelihood of converting to S-corporations. As dividend payout increases, so does the likelihood that the bank will convert. Moving from the 25th to the 75th percentile of DIVIDEND increases the probability of conversion by 14.1 percent, consistent with our hypothesis and prior literature. This result suggests that the elimination of double taxation is a significant tax benefit of conversion. Conversely, as corporate taxable income (TAXINC) increases, the likelihood of conversion decreases (consistent with Omer et al. [2000]), although the effect is smaller than for dividends.

---

[17] In comparison, the proportional chance criterion (C_{pro}), used in McKee et al. (1984) accurately predicts conversions 68.9 percent of the time. The proportional chance benchmark is $C_{PRO} = \alpha^2 + (1 - \alpha)^2$, where α is the observed proportion of S-converters and $(1 - \alpha)$ is the observed proportion of nonconverters. For our pooled sample of 1,272 banks that converted to S-corporations and 5,350 banks that did not convert to S-corporations by the end of 1999, the probability that the classification is S is 3.6 percent ($((1{,}272/6{,}622)^2)$). The corresponding probability for the nonconverters is 65.3 percent. The sum of these is 68.9 percent, which represents the chance probability of correct assignment. Overall, our model does only a slightly better job of predicting the organizational form choice (70 percent vs. 68.9 percent). However, our main focus is on identifying factors that *explain* banks' choices (rather than merely *predicting* their choices), and we rely on the sensitivity measures to interpret our results.

TABLE 2

Results of Logistic Regressions Explaining S-Corporation Conversion Decisions for 1,272 Banks That Convert from 1997 to 1999 and 5,350 Banks That Do Not Convert by 1999

$$S_t = \alpha + \beta_1\, DIVIDEND_{94\text{–}96} + \beta_2\, TAXINC_{94\text{–}96} + \beta_3\, BUILTIN_{t-1} + \beta_4\, LOSSCF_{t-1} + \beta_5\, AMT_{94\text{–}96} + \beta_6\, STATE_{t-1}$$
$$+ \beta_7\, GROWTH_{94\text{–}96} + \beta_8\, REGULATORY_{t-1} + \beta_9\, LEVERAGE_{94\text{–}96} + \beta_{10}\, SIZE_{94\text{–}96} + \beta_{11}\, AGE_{99}$$

	Variable	*Predicted Sign*	*Model 1*[a] *Coefficient (p-value)*	*Elasticity*[b]	*Sensitivity*[c] *(percentage change)*	*Model 2*[a] *Coefficient (p-value)*	*Sensitivity*[c] *(percentage change)*
	Intercept		−1.88 (<0.01)			−1.87 (<0.01)	
Tax	DIVIDEND	+	26.30 (<0.01)	0.14	14.1	—	—
	TAXINC	−	−8.91 (<0.01)	−0.11	−7.1	—	—
	BUILTIN	−	−243.10 (<0.01)	−0.27	−29.7	−239.10 (<0.01)	−30.1
	LOSSCF	−	−0.90 (<0.01)	—	−54.6	−0.79 (<0.01)	−49.8
	AMT	+	53.80 (<0.01)	0.11	16.8	58.20 (<0.01)	19.8
	STATE	+	0.13 (0.04)	—	11.4	—	—
	TAX BENEFIT	+	—	—	—	6.99 (<0.01)	39.3
Nontax	GROWTH	−	−0.19 (<0.01)	−0.05	−4.5	−0.19 (<0.01)	−4.5
	REGULATORY	−	−0.37 (0.01)	—	−27.1	−0.40 (0.01)	−28.6
	LEVERAGE	±	−0.95 (0.21)	−0.02	−1.9	−1.26 (0.09)	−2.4

(continued on next page)

TABLE 2 (continued)

	Variable	*Predicted Sign*	*Model 1*[a] *Coefficient (p-value)*	*Elasticity*[b]	*Sensitivity*[c] *(percentage change)*	*Model 2*[a] *Coefficient (p-value)*	*Sensitivity*[c] *(percentage change)*
Control	SIZE	±	−0.03	−0.29	−3.2	−0.04	−3.0
			(0.47)			(0.33)	
	AGE	±	0.24	0.82	24.6	0.25	29.6
			(<0.01)			(<0.01)	
Model Chi-Square (d.f. = 11)			294.21			267.58	
(p-value)			(<0.01)			(<0.01)	
% correctly predicted		Overall:	70.0			67.9	
		Banks that convert:	72.3			70.1	
		Nonconverting banks:	58.0			56.7	
Pseudo-R^2[d]			7.0%			6.4%	

---

[a] Model 1 includes separate components of the tax benefit: DIVIDEND, TAXINC, and STATE. Model 2 includes the combined TAX BENEFIT.

[b] The elasticity represents the responsiveness of the probability of converting to an S-corporation to small changes in the independent variables. We calculate elasticities for the continuous variables (with all variables set at their mean values) using the following as prescribed by LeClere (1992, 772):

$$\eta = \frac{X_k}{P_1}\frac{\partial P_1}{\partial X_k} = X_k \frac{e^{-(\alpha+X\beta)}}{(1 + e^{-(\alpha+X\beta)})} \beta_k.$$

[c] The sensitivity measures represent the change in the probability of converting to an S-corporation when each continuous variable changes from its 25th percentile value to its 75th percentile value and each categorical variable changes from 0 to 1 in turn. LeClere (1992, 772) prescribes this method.

[d] Psuedo-R^2 is calculated as follows: $\frac{R^2}{R^2_{max}} = \frac{1 - \left[\frac{L(0)}{L(\hat{\beta})}\right]^{2/n}}{1 - [L(0)]^{2/n}}$ where L(0) is the likelihood of the intercepts-only model, $L(\hat{\beta})$ is the likelihood of the specified model, and n is the sample size.

(*continued on next page*)

TABLE 2 (continued)

Variable Definitions:

S = 1 if bank converted to an S-corporation during 1997–1999, and 0 otherwise;

DIVIDEND = dividends scaled by assets, 1994–1996 average;

TAXINC = current tax provision divided by the highest corporate tax rate (35 percent), then scaled by assets, 1994–1996 average;

TAX BENEFIT = Federal and state tax benefit defined as the present value of 1994–1996 average dividends (assuming an 8 percent discount rate for a period of 10 years) × $(\tau_i + \tau_{is})$ + 1994–1996 average taxable income × $[(\tau_c - \tau_i) + (\tau_{cs} - \tau_{is})]$ where τ_c is the highest corporate federal tax rate, τ_i is the highest individual federal tax rate, τ_{cs} is the highest corporate state tax rate if a bank operates in a state that recognizes S-corporations, and 0 otherwise; and τ_{is} is the highest individual state tax rate if a bank operates in a state that recognizes S-corporations, and 0 otherwise;

BUILTIN = market value less book value of investment securities available for sale, scaled by average assets, measured in the year prior to conversion (1998 for nonconverting banks);

LOSSCF = 1 if the prior year's income tax provision is 0 or the prior year's income tax provision is negative and cannot be fully absorbed by the previous two years' income tax, and 0 otherwise; LOSSCF is measured in 1998 for nonconverting banks;

AMT = a proxy for the bank's alternative minimum tax, defined as income from tax-exempt leases and securities, scaled by assets, 1994–1996 average;

STATE = 1 if bank has a holding company *and* is located in a state without combined reporting *and* S-corporations are recognized for state purposes, and 0 otherwise, measured in the year prior to conversion (1998 for nonconverting banks);

GROWTH = annual percentage change in total assets, 1994–1996 average;

REGULATORY = 1 if bank is above 95th percentile of deferred tax assets to regulatory capital, *or* if writing off deferred tax assets prior to conversion causes bank to fall one or more regulatory capital categories, *or* to fall below the 8 percent regulatory capital minimum, and 0 otherwise, measured in the year prior to conversion (1998 for nonconverting banks);

LEVERAGE = nondeposit liabilities to assets, 1994–1996 average;

SIZE = natural log of assets (in thousands), 1994–1996 average; and

AGE = natural log of number of years since incorporation as of 1999.

As predicted, built-in gains on available-for-sale securities decrease the likelihood that a bank will convert to an S-corporation, consistent with the Omer et al. (2000) study that is based on firms in the natural resources industry. Our results are consistent with potential penalty taxes on built-in gains deterring conversion by eliminating banks' flexibility to defer the individual taxes on realized gains. The sensitivity measure indicates that this is a significant explanator of banks' conversion decisions—moving from the 25th to the 75th percentile of BUILTIN reduces conversion likelihood by almost 30 percent.

We find that banks with tax loss carryforwards (LOSSCF) are less likely to convert to S-corporations, as predicted. This result suggests that past losses that banks may only use to shelter future corporate income discourage conversion. However, neither Omer et al. (2000) nor Ayers et al. (1996) find that loss carryforwards are important in the conversion decision. The significance of loss carryforwards in our study may arise because our banking sample is more homogenous than other studies' samples that include industrial firms.

Converting banks can potentially avoid paying AMT. We find that banks' overall investment in tax-exempt securities (a proxy for AMT) is positively associated with conversion, which suggests that banks exposed to the AMT are more likely to convert to S-corporations.[18]

State taxes appear to influence conversion decisions. In contrast to Omer et al. (2000) but consistent with Plesko (1994), we find that banks that enjoy preferred state-tax treatment as S-corporations are more likely to convert. The lack of significance in Omer et al. (2000) likely stems from the concentration of their natural resource firms in states that do not recognize S-corporations. Our sample, in contrast, is more widely dispersed among states that do and do not recognize S-corporations.

Nontax Variables

Consistent with our prediction, the estimated coefficient on GROWTH is negative—banks with historically higher growth rates are less likely to convert to S-corporations. To the extent that historic growth proxies for future capital needs, our results suggest that restrictions on additional external funding reduce the likelihood that banks will convert to S-corporations. This result is consistent with Ayers et al. (1996), who find that S-corporation rules are more likely to restrict successful, high-growth firms than unsuccessful ones.

The coefficient on REGULATORY is negative—potential adverse regulatory ramifications of conversion dissuade banks from converting. The sensitivity measure indicates that REGULATORY is the most significant nontax explanator in our model. This represents a significant, and perhaps unintended, economic consequence of deferred tax accounting. Whereas other studies document that deferred tax accounting has economic consequences for publicly traded firms (Givoly and Hayn 1992; Ayers 1998), our results suggest that financial reporting rules affect private banks' organizational form decisions.[19]

---

[18] An alternative AMT measure leads to identical inferences. To satisfy pledging requirements, banks must pledge Treasury or (tax-exempt) municipal securities as collateral against certain deposits from the U.S. Treasury, municipalities, and Federal Reserve banks. A bank whose tax-exempt securities exceed its pledging requirements is not likely grappling with AMT; otherwise, it would divest the excess tax-exempt securities to avoid the additional tax AMT would impose. Our alternate AMT measure, tax-exempt securities in excess of pledging requirements, identifies banks *least* concerned with AMT exposure so these banks should be *less* likely to convert to S-corporations. Our results are consistent with that prediction. Taken together, our two AMT results suggest that when the alternative minimum tax applies, banks are more likely to reduce their exposure to the AMT by converting to S-corporations.

Contrary to prior research, we find that leverage is not associated with conversion decisions, perhaps because banks differ significantly from industrial firms included in previous studies. In particular, banks' leverage is not comparable to that of other firms because insured deposit liabilities figure so prominently in banks' capital structures.

Control Variables

Of the two control variables, only AGE is significantly associated with banks' conversion decisions: older, potentially more established banks are more likely to convert. This is consistent with younger, smaller banks hesitating to restrict their future access to capital (Ayers et al. 1996).[20] The coefficient on SIZE is not significant.

To ensure that extreme values do not drive our results, we winsorize our continuous variables at the 5th and 95th percentiles as well as at the 10th and the 90th percentiles. Estimating our model using these winsorized data sets leads to results similar to those reported in Table 2 (except AGE becomes insignificant). Additionally, we run the logistic regressions excluding any nonconverter bank that is larger than the largest bank that did convert. This reduces the sample by 37 banks but yields identical inferences.

Estimated Tax Benefit of S-Corporation Conversions

To quantify the magnitude of the potential tax benefit, we estimate the annual federal and state tax benefits from conversion using the following equation:

$$\begin{aligned}\text{Tax Benefit} = {} & \text{Dividends} \times (\tau_i + \tau_{is}) \\ & + \text{Taxable income} \times [(\tau_c - \tau_i) + (\tau_{cs} - \tau_{is})] \qquad (3)\end{aligned}$$

where τ_c is the highest corporate federal tax rate, τ_i is the highest individual federal tax rate, τ_{cs} is the highest corporate state tax rate if a bank operates in a state that recognizes S-corporations, and 0 otherwise; and τ_{is} is the highest individual state tax rate if a bank operates in a state that recognizes S-corporations, and 0 otherwise. We then compute the present value of the tax benefit assuming an 8 percent discount rate for a period of ten years. Model 2 of Table 2 presents the results of the logistic regression when we substitute this aggregate tax benefit variable for DIVIDEND, TAXINC, and STATE. As expected, we find that as the tax benefits increase, so does the likelihood that the bank will convert to an S-corporation. Moving from the 25th to the 75th percentile of TAX BENEFIT increases the probability of conversion by 39.3 percent, consistent with our hypothesis and prior literature. This result suggests that the tax benefit significantly influences banks' decisions to convert. All other variables retain their sign and significance levels with the exception of LEVERAGE, which is negative and marginally significant.

---

[19] One source of deferred tax assets is tax loss carryforwards. Because we measure REGULATORY with deferred tax assets, it is possible that the regulatory effect is confounded with the tax loss carryforward effect. However, untabulated frequency results show that only two of the 1,272 converters and 48 of the 5,350 nonconverters face adverse regulatory consequences of conversion *and* have net operating loss carryforwards. Given these low frequencies, we do not believe that the reduction in deferred tax assets is simply attributable to the loss of NOLs.

[20] While Omer et al. (2000) find that AGE is negatively related to conversion, their research design measures the marginal effects of AGE when tax rates changed around TRA86, whereas we measure the effect of AGE on banks making the organizational form choice for the first time, holding constant the difference between individual and corporate rates.

Although TAX BENEFIT has the expected positive effect on conversion, the explanatory power of Model 2 is lower than in Model 1. This is likely due to two reasons. First, aggregating DIVIDEND, TAXINC, and STATE into a single variable restricts the coefficients to be equal, whereas the unrestricted model suggests that avoiding double taxation on dividends plays a larger role than the higher personal tax rate on corporate earnings. Second, our estimation of the state tax portion of TAX BENEFIT introduces additional measurement error because we cannot estimate the additional state tax benefits that arise for banks that operate in states that disallow unitary reporting.

To provide further evidence comparing converters' and nonconverters' tax benefits, we present the estimated tax benefit for each group, in aggregate and by size quartiles in Table 3.[21] On average, we estimate that converters have a tax benefit of $1,479,310, compared to $934,470 for nonconverters. The tax benefit is largest for early converters ($1,728,120), is lower for late converters ($1,278,350), and is lower still for nonconverters ($934,470). Banks with more to gain converted immediately and those with smaller potential tax benefits took longer to convert, possibly because these banks needed additional time to plan the conversion.

Across all bank-size quartiles, we estimate that converters have larger mean and median tax benefits than nonconverters. Table 3 shows that a significant number of banks in the largest size quartile convert to S-corporations, indicating that some larger banks are willing to incur potentially high costs to eliminate the double tax burden.

V. EVIDENCE THAT BANKS MAKE STRATEGIC CHANGES BEFORE CONVERTING TO S-CORPORATIONS

In this section, we seek to explain why banks that converted did not uniformly do so in 1997, immediately after the tax-law change. First, we explore whether banks altered four disqualifying capital structure conditions in order to become S-corporations—public ownership, foreign shareholders, preferred shares, and more than 75 shareholders. Then we investigate whether banks managed three firm attributes to reduce conversion costs or increase conversion benefits—net operating losses, built-in gains, and dividends paid. To explore these issues, we analyze how the banks' capital structure, net operating losses, built-in gains, and dividends paid changed in the year before the bank converted. We compare the changes in these variables among three subsamples. That is, we compare early (1997) converters to late (1998 or 1999) converters, and we compare all converters to nonconverters (banks that had not converted by the end of 1999). Differences among these subsamples' changes in the period before conversion suggest that banks took strategic actions in anticipation of conversion.

Evidence That Banks Altered Disqualifying Capital Structure Conditions

Public Ownership

Panel A of Table 4 shows that no banks that were publicly traded in 1996 converted to S-corporations in 1997, presumably because delisting is a time-consuming process that could not be completed in 1996. In contrast, four banks that converted in 1998 and 1999 voluntarily delisted their publicly traded stock prior to converting into S-corporations. Press

---

[21] To compare converting and nonconverting banks of similar size, we eliminate 37 nonconverting banks that have 1994–1996 mean average assets larger than the largest 1994–1996 mean average assets of converting banks.

TABLE 3
Estimated Present Value of Federal and State Tax Benefits for Banks That Convert Compared to Banks That Do Not Convert by 1999[a]

(in $000)

	n	*Mean*	*Median*	*Standard Deviation*
Full Sample				
Banks that convert, 1997 to 1999	1,272	$1,479.31	$897.74	$2,524.48
Banks that do not convert by 1999	5,313	934.47*	293.48*	5,509.16
Banks that convert in 1997	575	1,728.12	1,073.96	2,899.37
Banks that convert in 1998 or 1999	697	1,278.35*	720.15*	2,148.31
1st Quartile—Smallest Banks				
Converters (Assets = $17,190)	293	383.28	330.85	288.66
Nonconverters (Assets = $16,034)	1,353	157.52*	80.89*	871.27
2nd Quartile				
Converters (Assets = $34,662)	383	780.41	680.90	546.70
Nonconverters (Assets = $34,940)	1,263	354.91*	229.31*	757.91
3rd Quartile				
Converters (Assets = $61,048)	347	1,480.26	1,329.08	909.01
Nonconverters (Assets = $62,193)	1,300	775.35*	512.92*	2,906.33
4th Quartile—Largest Banks				
Converters (Assets = $156,510)	249	3,842.70	2,909.60	4,813.37
Nonconverters (Assets = $176,989)	1,397	2,358.98*	1,085.92*	10,170.16

* Differences between converting and nonconverting banks are significant at a 0.10 level or less, based on two-sample t-tests and Wilcoxon signed-rank tests.

[a] The sample includes converting banks during 1997–1999, and banks no larger than the largest converter measured in 1999. This screen reduces the number of nonconverting banks from 5,350 to 5,313. We use this screen so we can compare the tax savings of similarly sized banks.

Variable Definition:

TAX BENEFIT = Federal and state tax benefit defined as the present value of 1994–1996 average dividends (assuming an 8 percent discount rate for a period of 10 years) × $(\tau_i + \tau_{is})$ + 1994–1996 average taxable income × $[(\tau_c - \tau_i) + (\tau_{cs} - \tau_{is})]$ where τ_c is the highest corporate federal tax rate, τ_i is the highest individual federal tax rate, τ_{cs} is the highest corporate state tax rate if a bank operates in a state that recognizes S-corporations, and 0 otherwise; and τ_{is} is the highest individual state tax rate if a bank operates in a state that recognizes S-corporations, and 0 otherwise.

releases for two of the delistings cited the banks' conversion to S-corporation as the reason for delisting. We conclude that for a very few banks, S-corporation benefits outweigh the benefits that accrue from access to public equity markets.

Foreign Shareholders

Tax laws preclude S-corporations from having foreign shareholders. Table 4, Panel B shows that no early converters had foreign shareholders, but 17 late-converter banks retired foreign-held shares to conform to the S-corporation proscription on foreign shareholders. Among nonconverters, about 3.2 percent had foreign shareholders and this did not change

TABLE 4
Disqualifying Capital Structure Conditions for Banks That Convert in 1997 (Early), Banks That Convert in 1998 or 1999 (Late), and Banks That Do Not Convert by 1999[a]

	Banks that Convert in 1997 (n = 543)	*Banks that Convert in 1998 or 1999 (n = 676)*	*Banks that Do not Convert by 1999 (n = 5,198)*
Panel A: Public Ownership[b]			
Number (%) of publicly traded banks in 1996	0	4 (0.59)	2 (0.04)
Number (%) of publicly traded banks in 1999	0	0	11 (0.21)*†
Panel B: Foreign Shareholders			
Number (%) of banks with foreign owners in 1996	0	17 (2.51)	164 (3.16)*†
Number (%) of banks with foreign owners in 1999	0	0	174 (3.35)*†
Panel C: Preferred Stock Activity			
Preferred stock in 1996 (in $000)	1.95	11.87	88.46*†
Preferred stock in 1999 (in $000)	0	0	125.68*†
% of banks with preferred stock issuing common stock in 1996	25.00	25.00	23.00
% of banks with preferred stock issuing common stock in year before bank converts (or in 1999 for nonconverters)	25.00	38.00*	20.00†
Panel D: Common Stock Activity			
% of banks retiring stock in 1996	3.50	1.87	1.79†
% change in stock retirement in year before bank converts	233.33	353.85*	−5.21*†
% of banks issuing common stock in 1996	1.75	1.72	7.68*†
% change in common stock issuance in year before bank converts (or in 1999 for nonconverters)	−8.85	−27.49*	−2.18†

---

* Difference from means of converting banks in 1997 is significant (at a 0.10 level or less).

† Difference from means of converting banks in 1998 or 1999 is significant (at a 0.10 level or less).

[a] To be included in this table, a bank must have data for all years from 1995 to 1999. This data requirement reduces the number of banks as compared to Tables 1 through 3.

[b] To be included in the public ownership statistics, a bank must have been private in 1997. The table reports the number of banks that were publicly traded before 1997.

appreciably from 1996 to 1999. Thus, for some converting banks, S-corporation benefits outweighed the cost of eliminating prohibited shareholders.

Preferred Stock

Table 4, Panel C shows that both early and late converters eliminated disallowed preferred stock prior to conversion. During the same time period, nonconverters increased preferred stock. Moreover, in the year prior to conversion, late converters with preferred stock issued common stock almost twice as often as nonconverters (38 percent compared to 20 percent). This suggests that late converters exchanged common shares for preferred shares prior to conversion, rather than retiring preferred shares outright.

Shareholders in Excess of Maximum

Banks with more than 75 shareholders cannot become S-corporations. We are unable to measure the number of shareholders directly. However, evidence of increased common-share *retirement* and decreased common-share *issuance* before conversion would be consistent with banks reducing the number of shareholders. In the year before converting to S-corporations, banks significantly increased the number of shares retired. Panel D of Table 4 shows that banks that converted in 1997 increased stock retirement by 233 percent, and banks that converted in 1998 or 1999 increased stock retirement by 353 percent. In contrast, banks that did not convert to S-corporations decreased share retirement by 5.21 percent.

Table 4, Panel D also shows that converters were less likely to issue common stock in 1996 than nonconverters (1.75 percent compared to 7.68 percent). In the year prior to conversion, stock issuance dropped by 8.85 percent for early converters and 27.5 percent for late converters. In contrast, stock issuance dropped by only 2.18 percent for nonconverters. Collectively, these results are consistent with converting banks retiring stock and curtailing new stock issuances in the year before conversion in order to meet the S-corporation rules for maximum number of shareholders.

Evidence That Banks Managed Net Operating Loss Carryforwards, Built-In Gains, and Dividends Paid to Increase Net Benefits of S-Corporation Conversion

Net Operating Loss Carryforwards

Banks that convert to S-corporations forgo benefits from net operating loss carryforwards. As expected, the incidence of net operating loss carryforwards among converting banks decreases in the year before conversion (see Table 5, Panel A). Among late converters, the number of banks with operating loss carryforwards decreased by 35.65 percent in the year before conversion compared to 1996. A similar pattern is not evident for nonconverters—the incidence of loss carryforwards increased by 9.60 percent from 1997 to 1998. This evidence is consistent with banks delaying their conversion until after they have taken advantage of the tax shield provided by their existing net operating loss carryforwards. An alternate explanation is that banks strategically reduce loss carryforwards in anticipation of S-conversion. Our results on realized gains (below) suggest that banks may have actively managed income to reduce net operating loss carryforwards.

Built-In (Unrealized) Security Gains

Because built-in (unrealized) gains on available-for-sale securities increase banks' costs of converting to S-corporations, we expect converters to reduce built-in gains before converting. We examine the annual changes in built-in (unrealized) gains over the two years before conversion. First, we calculate the level of unrealized gains (scaled by the size of banks' portfolios of securities available for sale) at the time of conversion (time t) as well as one and two years prior to conversion (time $t - 1$ and $t - 2$, respectively). Then we calculate year-over-year changes in unrealized gains. We expect that these changes will be negative in each of the two years immediately preceding a bank's conversion year for late converters. Early converters had little time to manage unrealized gains—there were only 11 days between the IRS ruling (Notice 97-5, 1997-1 CB 352, IRS 1996) and the end of the 1996 fiscal year (December 31 for all banks), when built-in gains would be calculated.

TABLE 5

Firm Attributes Associated with Costs and Benefits of a Conversion to S-Corporation, for Banks That Convert in 1997 (Early), Banks That Convert in 1998 or 1999 (Late), and Banks That Do Not Convert by 1999[a]

	Banks that Convert in 1997	*Banks that Convert in 1998 or 1999*	*Banks that Do not Convert by 1999*
Panel A: Net-Operating Loss Carryforwards	*(n = 543)*	*(n = 676)*	*(n = 5,198)*
% of banks with net-operating loss carryforwards in 1996	2.73	4.97*†	10.00*†‡
% change in net-operating loss carryforwards in year before bank converts (or in 1998 for nonconverters)	0.36	−35.65*†	9.60*†‡
Panel B: Gains and Losses on Securities	*(n = 500)*	*(n = 648)*	*(n = 4,856)*
Built-in gains on securities available for sale (AFS)			
$\Delta\text{BUILTIN}_{t-2}$	3.15*	6.45	−1.39*†
$\Delta\text{BUILTIN}_{t-1}$	−1.49*	3.91*†	3.96*†
$\Delta\text{BUILTIN}_{t-1} - \Delta\text{BUILTIN}_{t-2}$	−4.65*	−2.55	5.35*†‡
Realized security gains and losses			
$\Delta\text{RZD GL}_{t-2}$	−2.69	2.04*	−3.42‡
$\Delta\text{RZD GL}_{t-1}$	−0.52	27.25*†	5.86*†‡
$\Delta\text{RZD GL}_{t-1} - \Delta\text{RZD GL}_{t-2}$	2.18	25.20*†	9.38*
Panel C: Dividends Paid	*(n = 480)*	*(n = 580)*	*(n = 4,790)*
$\text{DIVIDENDS}_{1996}/\text{Pre-Tax Income}_{1996}$	42.34%*	36.3%*†	26.69%*†‡
$\Delta\text{DIVIDENDS}_{t-2}$	3.74%	8.27%*	8.48%*†
$\Delta\text{DIVIDENDS}_{t-1}$	11.42%*	0.92%†	5.53%*†‡
$\Delta\text{DIVIDENDS}_{t}$	78.20%*	81.23%*	14.21%*†‡
$\Delta\text{DIVIDENDS}_{t} - \Delta\text{DIVIDENDS}_{t-1}$	66.78*	80.31*†	8.68*†‡
$\text{DIVIDENDS}_{t}/\text{Pre-Tax Income}_{t}$	64.42%*	72.13%*	33.74%*†‡

---

* Variable mean is significantly different from 0 (at a 0.10 level or less).

† Difference from means of converting banks in 1997 is significant (at a 0.10 level or less).

‡ Difference from means of converting banks in 1998 or 1999 is significant (at a 0.10 level or less).

[a] To be included in this table, a bank must have data for the variables within each panel for all years from 1995 to 1999. This data requirement reduces the number of observations as compared to Tables 1 through 3. Nonconverters' variables are calculated in 1998.

Variable Definitions:

$\text{BUILTIN}_t = (\text{AFS Fair Value}_t - \text{AFS Book Value}_t)/\text{AFS Fair Value}_t$;

$\Delta\text{BUILTIN}_t = (\text{BUILTIN}_t - \text{BUILTIN}_{t-1})/(|\text{BUILTIN}_{t-1}|)$;

RZD GL_t = realized security gains and losses reported on bank's income statement for year t;

$\Delta\text{RZD GL}_t = (\text{RZD GL}_t - \text{RZD GL}_{t-1})/(|\text{RZD GL}_{t-1}|)$;

DIVIDENDS_t = cash dividends paid on common stock during year t; and

$\Delta\text{DIVIDENDS}_t = (\text{DIVIDENDS}_t - \text{DIVIDENDS}_{t-1})/\text{DIVIDENDS}_{t-1}$. If DIVIDENDS_{t-1} is equal to 0, then we scale by DIVIDENDS_{t-2}.

Thus, we expect only the t − 1 change in unrealized gains will be negative for the early converters.

Consistent with expectations, Table 5, Panel B shows that early converters' built-in gains decreased in the year before conversion. Specifically, gains were 1.49 percent lower at conversion than they were one year earlier. Moreover, this decrease in unrealized gains is material relative to the increase in unrealized gains these early converters experienced in the preceding year (the t − 1 change in unrealized gains is 4.65 percentage points lower than the t − 2 change). The decline in unrealized gains is consistent with early converters managing built-in gains prior to conversion. However, late converters' built-in gains increased in both years prior to conversion—built-in gains increased 6.45 percent in t − 2 and 3.91 percent in t − 1, although only the t − 1 change is statistically significant. We next use time-series and cross-sectional tests to investigate whether overall market conditions could have driven this puzzling result in the 1997–1999 period.

In the ordinary course of business, a bank realizes gains and losses as it realigns its securities portfolios for liquidity, profit-taking, and other reasons. Absent strategic intervention, *realized* portfolio gains and losses ought to change proportionately with the *unrealized* gains and losses left in the portfolio. That is, realized gains that exceed unrealized gains potentially indicate selective gain recognition (Beatty 1995). Panel B of Table 5 reports the changes in realized gains in banks' securities portfolios one and two years before conversion. Late converters' realized gains in the t − 2 period increased by roughly the same magnitude as unrealized gains during the same period (2.04 percent and 6.45 percent respectively). However, in the t − 1 period, late converters' realized gains increased by 27.25 percent, or roughly ten times the increase in unrealized gains (3.91 percent). This time-series comparison suggests that 1998 and 1999 converters strategically realized some *but not all* of the gains in their available-for-sale securities portfolios.

Cross-sectional tests corroborate this conclusion; banks that did not convert by 1999 appear to have behaved differently than late converters. Table 5, Panel B shows that nonconverters' realized and unrealized gains changed by roughly the same magnitude in the t − 2 and the t − 1 periods. Taken together, this evidence suggests that prior to conversion, converters strategically reduced built-in (i.e., unrealized) gains by selectively selling available-for-sale securities with larger gains and leaving securities with smaller gains in the portfolio. Nonetheless, late converters apparently were not able to completely counter the large market upswings during the 1998–1999 period, and built-in gains remained in late converters' portfolios.

Dividend Payout

A bank that intends to convert to an S-corporation could avoid double taxation on dividends in the last year or years as a C-corporation by temporarily decreasing dividends. After S-conversion, the bank could pay the postponed dividends to make the shareholders whole. Contrary to our expectations, we do not find that converting banks significantly reduced dividends in the years before conversion. Instead, Table 5, Panel C reports that average dividends increased in both the t − 2 and the t − 1 periods. However, when we compare dividend increases over time and between converters and nonconverters, we conclude that late converters postponed dividend increases before converting to S-corporations. Specifically, in the year before converting, late converters' dividend increases were significantly smaller than those for the prior year (average t − 2 increase of 8.27 percent compared to average t − 1 increase of 0.92 percent). Moreover, relative to nonconverters, late converters increased dividends by a significantly smaller amount in the year before conversions: 5.53 percent increase for nonconverters and 0.92 percent increase for late converters. This provides some evidence that banks that converted to S-corporations in 1998 and 1999 had lower dividend growth relative to banks that did not convert. This strategic manipulation of dividends is consistent with converters' incentives to postpone dividends in order to increase conversion benefits. We do not find any evidence (either time-series or cross-sectional) that banks that converted to S-corporations in 1997 managed dividends. Again, this is likely due to the short time interval (11 days) between the law enactment and the end of banks' fiscal year on December 31, 1996.

Summary

Taken together, our findings suggest that certain capital structure regulations contribute to late conversion. We document changes in capital structure that are consistent with banks overcoming disqualifying conditions prior to conversion. We also observe significant, predictable changes in converting banks' attributes consistent with banks increasing the net benefits of conversion (e.g., accelerating the use of net operating loss carryforwards, realizing gains on sales of securities, and reducing dividend increases). The fact that we do not observe similar patterns of behavior among nonconverters suggests that banks that converted to S-corporations behaved strategically.

VI. CONCLUSIONS

The natural experimental setting created by the Small Business Job Protection Act of 1996 and the unique database of privately held banks allow us to conduct more powerful tests of organizational form choice. We find that tax considerations are associated with banks' choice of organizational form. We conclude that banks making S-corporation conversion decisions consider both tax benefits from conversion and tax costs associated with the conversion. Results of our logistic model show that net tax savings is a significant factor in banks' decisions to convert. We also document that the state in which banks conduct business affects their conversion choices, which suggests that state tax regimes can counteract the incentives provided by federal statutes.

Nontax considerations such as future capital needs and firm age also appear to influence banks' conversion decisions. We identify a previously undocumented regulatory cost of SFAS No. 109—the elimination of deferred tax assets when a bank converts to a nontaxable entity such as an S-corporation reduces conversion likelihood. Specifically, banks whose regulatory capital would be most seriously affected by the elimination of deferred tax assets are about 39 percent less likely to convert. A limitation of our study is that our findings concerning the regulatory cost of conversion may not generalize to nonregulated industries.

Results of an exploratory analysis in the two-year period before banks convert to S-corporations are consistent with banks taking strategic actions to increase the benefits and decrease the costs associated with conversion. We find evidence that banks retired preferred stock, eliminated foreign shareholdings, and reduced tax loss carryforwards. Moreover, converting banks selectively realized security gains before conversion—*realized* gains increased by 27.25 percent, *unrealized* gains decreased by 3.91 percent. Banks that converted to S-corporations in 1998 and 1999 slowed dividend increases relative to historical increases and relative to nonconverting banks' increases. Taken together, these results suggest that banks undertook costly restructurings and managed operating transactions in order to reap S-corporation benefits.

By the end of 1999, only 19.2 percent of banks had converted to S-corporations—far fewer than for the general population of firms, in which 53.4 percent are S-corporations (Wittman 2001). Legislators and financial analysts expected that more banks would convert to S-corporations (U.S. Congress 1998). Although our model may have limited *predictive* ability, our findings suggest specific factors that *explain* banks' conversion decisions and that future legislation might address (e.g., eliminating the double tax on dividends, or relaxing built-in gains rules). For example, our results suggest that current congressional proposals to increase the maximum allowable number of shareholders from 75 to 150, may be effective in encouraging S-corporation conversion. In addition, the interplay between regulatory capital rules and deferred tax accounting standards (SFAS No. 109) was seemingly unanticipated, and our findings may inform future bank regulations that rigidly apply financial accounting standards. Finally, a recent GAO report discusses bank regulators' concerns about how proposals to further liberalize S-corporation rules could impair community bank safety, soundness, and liquidity (U.S. GAO 2000). Our results may help bank regulators estimate regulatory effects of S-corporations under existing and prospective tax laws.

REFERENCES

Amihud, Y., B. Lev, and N. Travlos. 1990. Corporate control and the choice of investment financing: The case of corporate acquisitions. *Journal of Finance* 45 (2): 603–616.

Ayers, B., C. B. Cloyd, and J. Robinson. 1996. Organizational form and taxes: An empirical analysis of small businesses. *The Journal of the American Taxation Association* 18 (Supplement): 49–67.

———. 1998. Deferred tax accounting under SFAS No. 109: An empirical investigation of its incremental value-relevance relative to APB No. 11. *The Accounting Review* 73 (2): 195–212.

Beatty, A. 1995. The effects of fair value accounting on investment portfolio management: How fair is it? *Federal Reserve Bank of St. Louis Review* 77 (1): 25–39.

———, and D. Harris. 2001a. The impact of explicit and implicit state taxation of U.S. government obligations on the structure of banks' investment and financing portfolios. *The Journal of the American Taxation Association* 23 (2): 1–19.

———, and ———. 2001b. Intra-group, interstate strategic income management for tax, financial reporting, and regulatory purposes. *The Accounting Review* 76 (4): 515–536.

Cloyd, C. B., S. Limberg, and J. Robinson. 1997. The impact of federal taxes on the use of debt by closely held corporations. *National Tax Journal* 50 (2): 261–277.

Damodaran, A., J. Kose, and C. Liu. 1997. The determinants of organizational form changes: Evidence and implications from real estate. *Journal of Financial Economics* 45 (2): 169–192.

Erickson, M., and S. Wang. 2000. The effect of transaction structure on price: Evidence from subsidiary sales. *Journal of Accounting and Economics* 30 (1): 59–97.

Financial Accounting Standards Board (FASB). 1992. *Accounting for Income Taxes.* Statement of Financial Accounting Standards No. 109. Norwalk, CT: FASB.

Gentry, W. 1994. Taxes, financial decisions, and organizational form: Evidence from publicly traded partnerships. *Journal of Public Economics* 53 (2): 223–244.

Givoly, D., and C. Hayn. 1992. The valuation of the deferred tax liability: Evidence from the stock market. *The Accounting Review* 67 (2): 394–410.

Guenther, D. 1992. Taxes and organizational form: A comparison of corporations and master limited partnerships. *The Accounting Review* 67 (1): 17–45.

Internal Revenue Service (IRS). 1996. *Cumulative Bulletin Notice 97–5, 1997–1 CB 352.* Washington, D.C.: Government Printing Office.

Klassen, K. 1997. The impact of inside ownership concentration on the tradeoff between financial and tax reporting. *The Accounting Review* 72 (3): 455–474.

LeClere, M. 1992. The interpretation of coefficients in models with qualitative dependent variables. *Decision Sciences* 23 (3): 770–776.

McKee, A., Jr., T. Bell, and J. Boatsman. 1984. Management preferences over accounting standards: A replication and additional tests. *The Accounting Review* 59 (4): 647–659.

Omer, T., and W. Terando. 1999. The effect of risk and tax differences on corporate and limited partnership capital structure. *National Tax Journal* 52 (4): 699–715.

———, G. Plesko, and M. Shelley. 2000. The influence of tax costs on organizational choice in the natural resource industry. *The Journal of the American Taxation Association* 22 (1): 38–55.

Petroni, K., and D. Shackelford. 1995. Taxation, regulation, and the organizational structure of property-casualty insurers. *Journal of Accounting and Economics* 20 (3): 229–253.

Plesko, G. 1994. Corporate taxation and the financial characteristics of firms. *Public Finance Quarterly* 22 (3): 311–334.

———. 1995. The role of taxes in organizational choice: S conversions after the Tax Reform Act of 1986. Working paper, Massachusetts Institute of Technology, Cambridge, MA.

Sawyers, R., and M. Beasley. 1998. The impact of state taxation on the investment portfolio of banks. *The Journal of the American Taxation Association* 20 (Supplement): 1–14.

Shaw, W., and H. Wier. 1993. Organizational form choice and the valuation of oil and gas producers. *The Accounting Review* 68 (3): 657–667.

Terando, W., and T. Omer. 1993. Corporate characteristics associated with master limited partnership formation. *The Journal of the American Taxation Association* 15 (1): 23–45.

U.S. Congress. 1986. *Tax Reform Act of 1986.* Public Law No. 99–514. Washington, D.C.: Government Printing Office.

———. 1996. *Small Business Job Protection Act of 1996.* Public Law No. 104–188. Washington, D.C.: Government Printing Office.

———. 1998. House Banking Committee Report 11/24/98. Washington, D.C.: Government Printing Office.

U.S. General Accounting Office (GAO). 1995. *Tax Policy: Experience with the Corporate Alternative Minimum Tax.* April. Washington D.C.: Government Printing Office.

———. 2000. *Banking Taxation: Implications of Proposed Revisions Governing S-Corporations on Community Banks.* June. Washington D.C.: Government Printing Office.

Wittman, S. 2001. S-corporation returns, 1998. *Statistics of Income Bulletin* 20 (4): 46–101.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 327–351

Residual Income Risk, Intrinsic Values, and Share Prices

Stephen P. Baginski
University of Georgia
James M. Wahlen
Indiana University

ABSTRACT: Empirical accounting research provides surprisingly little evidence on whether accounting earnings numbers capture cross-sectional differences in risk that are associated with cross-sectional differences in share prices. We address two questions regarding the risk-relevance of accounting numbers: (1) Are accounting-related risk measures (i.e., the systematic risk and total volatility in a firm's time-series of residual return on equity) associated with the market's assessment and pricing of equity risk? (2) If so, then are these accounting-related risk measures incrementally associated with the market's assessment and pricing of equity risk beyond other observable factors, such as those in the Fama and French (1992) three-factor model?

We develop an accounting-fundamentals-based measure of the market's pricing of risk—the difference between actual share price and a residual income valuation model estimate of share value using risk-free rates of return. Our results show that both systematic risk and total volatility in residual return on equity partially explain this *pricing differential*, and that the explanatory power of total volatility is incremental to the Fama and French (1992) factors—market beta, firm size, and the market-to-book ratio.

Keywords: *residual income; accounting risk measures; implied risk premium; equity valuation.*

We appreciate research assistance and comments from M. Kimbrough, and comments from two anonymous referees, G. Clinch (the associate editor), D. Barrett, B. Beaver, W. Blacconiere, P. Easton, I. Jansen, B. Jorgensen, L. Maines, G. Miller, J. Ohlson, S. Penman, J. Salamon, M. Wallmeier, and L. Vincent; participants at the Financial Economics and Accounting Conference at the University of Michigan, the First Annual Winter Accounting Research Conference at the University of Utah, the Third Annual AAA Globalization Conference in Berlin, the 2001 Stanford Summer Camp, and the 2002 AAA Midwest meetings; and workshop participants at the following: Arizona State University, Indiana University, University of Missouri, The Ohio State University, and Washington University in St. Louis. We also gratefully acknowledge the contribution of I/B/E/S International Inc. for providing earnings per share forecast data. These data have been provided as part of a broad academic program to encourage earnings expectations research. Professor Wahlen gratefully acknowledges financial support from the Ford Motor Company Teaching Fellowship.

Submitted January 2001
Accepted September 2002

Data Availability: *Data are commercially available from the sources listed in the text.*

I. INTRODUCTION

Fundamental valuation of equity shares requires estimation of expected future payoffs and the risk inherent in those payoffs. Existing research on the usefulness of accounting earnings numbers has devoted far more attention to their role as payoff-relevant information than to their role as risk-relevant information. One exception is the seminal Beaver et al. (1970) study, which shows that accounting-based risk measures are positively associated with market model beta, but which does not examine whether accounting-based risk measures explain share prices or returns. Thus far, the empirical accounting research literature has surprisingly little to say about whether accounting earnings numbers capture cross-sectional differences in risk that are associated with cross-sectional differences in share prices. In this study, we investigate the risk-relevance of accounting numbers by addressing the question: Are accounting earnings-based risk measures associated with the capital market's assessment and pricing of firm risk? The answer to this question will inform capital markets participants, as well as accounting researchers and teachers, about the fundamental usefulness of accounting earnings numbers in assessing and pricing risk. The findings will also contribute to a better understanding of how to specify accounting earnings-based valuation models, and how to use them in settings in which market-based risk measures (e.g., market model beta) are not available.

We also address a second question: Are accounting earnings-based risk measures incrementally associated with the market's assessment and pricing of equity risk beyond other observable risk factors, such as the three factors in the Fama and French (1992) model (market model beta, firm size, and book-to-market ratios)? Research by Fama and French (1992) and others shows that the single factor capital asset pricing model may be incomplete because *ad hoc* factors outside of the model (including factors based on accounting numbers, such as the book-to-market ratio) appear to explain stock returns. Our investigation contributes evidence on whether accounting earnings-based risk measures capture elements of priced risk that traditional measures of equity risk (e.g., market model beta) or factors identified by more recent *ad hoc* approaches to risk (e.g., Fama and French 1992) do not capture.

Traditional theory on the role of accounting numbers in valuation, such as the residual income valuation models (e.g., Ohlson 1995; Feltham and Ohlson 1995), simplify the role of risk by assuming that investors are risk neutral and discount rates are nonstochastic and flat. More recently, Feltham and Ohlson (1999) point out that equity values should price as fundamental risk the nondiversifiable variability inherent in expected future residual income. Feltham and Ohlson (1999) demonstrate analytically that (at least in principle) one can incorporate risk in residual income valuation by reducing expected future abnormal earnings to certainty equivalents based on investors' risk aversion across all possible events and dates.[1] In this demonstration, Feltham and Ohlson (1999) measure abnormal earnings as earnings less a charge for the cost of equity capital, basing the charge on the book value of equity and the term structure of risk-free interest rates at the time of valuation. The pricing of risk therefore depends on the appropriate set of event-date-contingent prices for future abnormal earnings measured as certainty equivalents. The Feltham and Ohlson (1999) demonstration is silent, however, on how investors and empirical researchers should develop this complete set of event-date-contingent prices.

---

[1] We use the terms *residual income* and *abnormal earnings* interchangeably.

In the absence of implementable theoretical guidance, empirical applications of residual income valuation models have incorporated risk into valuation by adding an *ad hoc* risk premium to a risk-free discount rate, with results that are understandably sensitive to the risk premium assumption (e.g., Bernard 1994, 1995; Francis et al. 2000, 2001; Dechow et al. 1999). Other recent studies invert the residual income valuation model to estimate the *ex ante* risk premia implicit in discount rates, with mixed results (e.g., Claus and Thomas 2001; Gebhardt et al. 2001; Easton et al. 2000; Botosan and Plumlee 2001, 2002). All of these prior studies use observed share values or stock returns to estimate the risk premia and expected rates of return required by valuation models. Using observed share values or stock returns to assess risk introduces a degree of circularity into valuation.

We develop a more direct approach that uses accounting numbers to assess firm risk and share values in a residual income valuation context. First, we develop an accounting-based measure of the discount for risk inherent in share prices. We estimate *risk-free value* based on the residual income model, analysts' forecasts of earnings, and prevailing risk-free rates of return. We then calculate the *price differential*—the risk-free value estimate minus share price. Conceptually, the price differential is a simple yet theoretically defensible measure of the discount for risk implicit in share price. This measure depends only on analysts' expectations of earnings, the residual income valuation model, time value of money at prevailing risk-free rates, and share price. The price differential is a potentially appealing firm-specific measure of the cost of risk because it does *not* depend on any functional form of expected returns, or on explicit parameter estimates of market risk measures (i.e., betas) or risk premia. As expected, price differentials are positive for nearly all firm-years because risk-free values ignore the discounts for risk in share prices. As a practical matter, our estimates of price differentials are highly positively correlated with estimates of expected rates of return implicit in share prices, but are simpler to compute.

We next measure systematic risk and total volatility inherent in the time-series of abnormal return on equity (abnormal ROE) computed as return on book value of equity minus the prevailing risk-free interest rate, thereby isolating the measurement of abnormal ROE from the measurement of risk.[2] We measure systematic risk in residual income as abnormal ROE beta estimated using a regression of each firm's abnormal ROE on a sample-wide index of abnormal ROE, analogous to the accounting earnings beta in Beaver et al. (1970).[3] We also examine the total volatility in residual income using the standard deviation of abnormal ROE over time, because prior research has shown that total volatility in earnings is a proxy for risk (Beaver et al. 1970).

We test whether systematic risk and total volatility in residual income explain the cross-section of price differentials. If price differentials capture the fundamental discount for risk in share prices, and if abnormal ROE beta and/or the standard deviation of abnormal ROE are reliable surrogates for priced risk factors, then we expect price differentials to increase with residual income risk.

Consistent with our hypothesis, results from portfolio and univariate regression tests show that abnormal ROE beta and total volatility in abnormal ROE are significantly positively related to price differentials. The portfolio test results indicate that such differentials can be economically large relative to share prices. However, although abnormal ROE beta is positively and significantly associated with price differentials in most sample years, it provides relatively little explanatory power for price differentials (often less than 1 percent). In contrast, total volatility in abnormal ROE is more strongly and positively related with price differentials in nearly all sample years. Overall, the results validate that capital markets participants and researchers can use accounting numbers together with the residual income valuation model to measure firm risk and the discount for risk implicit in share price. The findings suggest that capital markets participants and researchers can implement residual income valuation by directly assessing risk inherent in residual income, avoiding the circularity inherent in stock-market-based risk measures. Our evidence also suggests that when the data necessary to compute market-based risk measures (e.g., the three Fama and French [1992] factors) are not available (e.g., privately held firms or initial public offerings), total volatility in abnormal ROE is a useful risk proxy in valuation.

---

[2] Our use of risk-free rates to determine abnormal earnings is consistent with Feltham and Ohlson (1999).

[3] Like an earnings-based beta, abnormal ROE beta should measure time-series covariation between earnings and market-wide risk factors. In addition, abnormal ROE beta should also measure time-series covariation between market-wide risk factors and book values and dividends.

In multiple regressions of price differentials on risk factors, we find that the capital markets price abnormal ROE beta incrementally to the three factors in the Fama and French (1992) model and to controls for potential measurement error in price differentials (long-run earnings growth and analysts' forecast errors), but only in three of nine sample years. On the other hand, we find the capital markets price total volatility in abnormal ROE incrementally to the three factors in the Fama and French (1992) model and the measurement error controls in six of nine sample years, and total volatility in abnormal ROE provides the most explanatory power of *all* the risk measures for price differentials. Despite the fact that we do not measure total volatility in abnormal ROE using explicit systematic risk factors, it nevertheless captures the market's pricing of systematic risk factors, including risk factors beyond the Fama and French (1992) three-factor model.

In supplemental analyses, we compare the share price implications of our two measures of residual income risk against analogous risk measures based on reported net income. Not surprisingly, we find that the risk measures based on either reported income or residual income serve equally well as risk proxies for purposes of residual income valuation. Residual income risk measures are highly positively correlated with reported income risk measures, and both provide comparable explanatory power for price differentials.

Our measure of price differential is a conceptually sound, computationally simple, and empirically valid accounting-based measure of the discount for risk implicit in share price. We find that price differentials are related to our two measures of residual income risk, market model beta, and the other two Fama and French (1992) risk factors. The price differential measure of the discount in share prices for risk may be useful in future empirical accounting research, and in teaching courses in financial statement analysis and valuation.

We have organized the paper as follows. In Section II, we motivate our hypothesis by describing the accounting earnings valuation model, the role of residual income risk in the model, and the existing empirical evidence on risk from applications of the model. In Section III, we describe our empirical methodology, including our two measures of residual income risk and our measure of price differentials. We describe the results in Section IV, and offer concluding remarks and implications in Section V.

II. THE ACCOUNTING EARNINGS VALUATION MODEL AND RISK

Classical valuation theory describes *share values* as the present value of all expected future dividends to investors from the outcomes of the firm's operating and investing decisions, priced to reflect nondiversifiable risk. Research in finance and accounting commonly uses proxies for expected future dividends in assessing risk and estimating share values. For example, researchers typically use stock returns as proxies for changes in market expectations of future dividends in the assessment of nondiversifiable risk, and earnings and cash flows as proxies for dividends as the relevant payoffs to shares.

Preinreich (1938), Edwards and Bell (1961), Peasnell (1982), Ohlson (1995), and others show that, as long as forecasts of earnings, book values, and dividends follow clean surplus accounting, the dividend valuation model is equivalent to the residual income valuation model:

$$V_t = B_t + \sum_{i=1}^{\infty} \frac{E_t[X_{t+i} - rB_{t+i-1}]}{(1 + r)^i} \tag{1}$$

where V_t denotes share value at time t, B_t denotes accounting book value of equity per share at time t, E_t represents the expectations operator as of time t, X_{t+i} represents earnings per share in period t + i, and r denotes the required rate of return. This formulation shifts the valuation focus away from wealth distribution (dividends) to wealth in place (book values) and wealth creation (future abnormal earnings.)

Ohlson (1995) and Feltham and Ohlson (1995) assume risk neutrality and homogeneous expectations; thus r is a nonstochastic and flat risk-free rate of return. Feltham and Ohlson (1999) illustrate a general approach to incorporate risk in residual income valuation by reducing the distribution of expected future abnormal earnings to certainty equivalents, which are priced based on event-date-contingent prices that span all conceivable future abnormal earnings outcomes. Such certainty equivalent adjustments include discounts for risk insofar as the event-date-contingent prices depend on the way abnormal earnings covary with economy-wide risk factors. Feltham and Ohlson (1999) also show that, under a certainty equivalent approach, abnormal earnings should include a charge for equity capital based on the spot risk-free rates implicit in the (not necessarily flat) term structure of interest rates at the time of valuation.[4] Empirical tests of such a formulation, however, require that the researcher specify a utility function that captures investors' risk aversion in order to price all possible date- and time-contingent abnormal earnings outcomes, a task that has so far defied empirical application.

Ohlson (1995, 680) notes that one way to incorporate risk in empirical applications of the model is to replace the risk-free rate with a risk-adjusted expected return. Accordingly, most prior empirical studies of the properties of residual income model value estimates have assumed cross-sectionally constant discount rates based on prevailing risk-free rates plus an *ex post* market risk premium estimate, usually on the order of 6 or 7 percent (e.g., Bernard 1994, 1995; Francis et al. 2000, 2001; Dechow et al. 1999). Ohlson (1995) observes that this approach is practical, but that it lacks theoretical foundation because it is silent about the source of the risk. Risk adjustments should depend on the nondiversifiable risk inherent in future abnormal earnings (or equivalently, future earnings, dividends, and book values).

A number of recent studies invert the residual income valuation model to infer risk premia and discount rates conditional on share prices, book values, and analysts' earnings expectations. Claus and Thomas (2001) and Gebhardt et al. (2001) obtain results that suggest that discount rates and risk premia are lower than those implied by historical data on realized returns. On the other hand, Botosan and Plumlee (2001, 2002) obtain results indicating implicit discount rates that are high relative to historical realized returns. Easton et al. (2000) point out that implied discount rates estimated by inverting the residual income valuation model depend on the assumed rate of growth in terminal-year residual income. If terminal-year growth assumptions are too low, then the implied discount rate estimates will also be too low, and vice versa. Therefore, Easton et al. (2000) invert the residual income model and estimate simultaneously the implicit discount rates and terminal-year residual income growth rates. Their implicit discount rate estimates are consistent with historical stock returns, averaging 13 percent over their 1981–1998 sample period. They also find that implicit growth rates in terminal-year residual income are much higher than assumed in prior research, averaging 10 percent over their sample period.

---

[4] These results require no risk-free arbitrage in book values and future abnormal earnings, as well as the clean surplus accounting relation.

These *ex post* and *ex ante* approaches to estimate discount rates and share values are circular, however, relying on observed share prices to infer the discount rates required to value shares. We develop a more direct approach based on the fundamental notion of using accounting numbers to measure risk *and* to determine share values. We also extend Beaver et al. (1970) by directly examining how residual income-based risk measures are associated with share prices. As noted earlier, the residual income model shifts the focus of valuation to wealth creation; we therefore focus on the risk inherent in wealth creation (residual income).

To measure the discount for risk implicit in share prices, we first compute *risk-free value* based on the residual income model using book values of equity, analysts' consensus expectations of future earnings, and prevailing risk-free rates of return. We then compute the *price differential* as the risk-free value estimate minus share price. The price differential is the market's discount for risk because it measures the difference between share value in a hypothetical risk-neutral market and observed share price. This measure depends only on analysts' expectations of earnings, the residual income valuation model, time value of money at risk-free rates, and share price.

We predict that price differentials increase with the risk inherent in residual income. We therefore develop and test two measures of residual income risk. First, the pricing of abnormal earnings should depend on the way abnormal earnings covary with systematic risk factors (Feltham and Ohlson 1999). Our first estimate of firm-specific systematic risk is an abnormal ROE beta, which we obtain by regressing abnormal ROE on a sample-wide index of abnormal ROE. Our measure of abnormal ROE beta is analogous to the earnings beta in Beaver et al. (1970).

Second, because other single-factor models of risk are incomplete (e.g., the single-factor capital asset pricing model), and because prior empirical work has shown that total variability in net income is a surrogate for risk (e.g., Beaver et al. 1970), our second estimate of firm-specific systematic risk is the total variability in residual income, measured as the standard deviation in the time-series of abnormal ROE. A limitation of this volatility-based risk measure is that it represents total rather than nondiversifiable risk. We nevertheless investigate this risk measure for practical reasons. As Beaver et al. (1970) note, estimates of systematic risk (whether based on stock returns or earnings) suffer from estimation error. Our tests should indicate whether the standard deviation in abnormal ROE, which is simple to estimate and requires less data than covariation-based measures of risk, is sufficient (i.e., does not lose information) for empirical applications of the residual income model.

If the capital markets price residual income risk and if abnormal ROE beta and the standard deviation of abnormal ROE capture such risk, then price differentials should increase across firms as residual income risk increases, all else held constant. More formally, we test the following hypothesis (stated in alternative form):

H1: Price differentials are increasing in abnormal ROE beta and the standard deviation in abnormal ROE, *ceteris paribus.*

The capital asset pricing model uses the covariation of stock returns with nondiversifiable, economy-wide risk factors to assess risk. Risk assessment based on covariation between systematic risk factors and either stock returns or accounting-earnings-based returns (i.e., abnormal ROE) are parallel in principle, but differ significantly in implementation. For example, our measures of abnormal ROE beta and total volatility in abnormal ROE lack precision, in part because firms report earnings infrequently (i.e., quarterly and annually). This lack of precision reduces the power of our tests of the pricing implications of accounting-based risk measures.

An even more serious concern is that residual income-based risk measures are imprecise because they are determined by the information that the accounting process recognizes in earnings and book values. By contrast, risk measures based on stock returns have an information advantage, insofar as share prices impound all value-relevant information. In addition, residual income risk measures may suffer from measurement error if the accrual accounting process imperfectly matches economy-wide risk factors with reported profits at the firm level. This could occur for any number of reasons, including conservatism (which, in some circumstances, delays recognition of good news), arbitrary cost allocation processes (e.g., straight-line depreciation of long-lived tangible assets or goodwill amortization), and earnings management.

On the other hand, accrual accounting may have certain information advantages for measuring risk. The accounting system aggregates the effects of transactions over time into periodic measures of quarterly or annual performance, which capture the extent to which the effects of risk are (or are not) diversified over time within each accounting period.[5] Earnings numbers measure the aggregate results of the firm's operating and investing activities each period, so the time-series of earnings should reflect the fundamental, long-run risk and volatility inherent in the firm's operations and investments. In addition, accrual accounting filters the effects of certain types of noise that influence stock-return-based measures of risk, stemming from excess volatility in returns (e.g., for *glamour stocks*), or temporary movements of prices away from fundamentals (e.g., noise trading, excessive exuberance, or the stock bubble in the late 1990s.)

Ultimately, the question of whether residual-income-based risk measures are economically meaningful measures of risk must be settled with empirical data. Therefore, we compare the pricing implications of residual income risk measures to the pricing implications of systematic risk in stock returns. We also test whether the capital markets price our two measures of residual income risk incrementally to the three Fama and French (1992) risk factors: market model beta, firm size, and book-to-market ratios.

III. EMPIRICAL METHODOLOGY

In this section we describe our sample selection and data collection. We also describe our key variables and discuss statistics that describe our sample data.

Sample Selection and Data Collection

Our sample of firms is the intersection of the I/B/E/S, CRSP, and Compustat Primary, Supplementary, Tertiary, and Research databases meeting the following requirements, applied yearly from 1990 to 1998:

1. From I/B/E/S (collected each April), consensus analysts' forecasts of one-year-ahead and two-year-ahead earnings per share and three- to five-year earnings growth, actual earnings per share, April 1 price, and dividends; in addition, consensus forecasts of two-year-ahead earnings per share must be positive;

---

[5] For example, suppose a firm exposes its operations to input and output price movements on a daily or weekly basis. That firm's quarterly and annual reports would report the aggregated (i.e., diversified over time) exposure to such risks.

2. From Compustat, a complete set of book values of common equity, annual earnings per share, and numbers of shares outstanding for the ten consecutive years prior to each valuation date; and
3. From CRSP, sufficient data to compute firm-specific betas from five years of monthly returns preceding each valuation date.

The sample meeting these requirements ranges from 665 firms in 1990 to 1,073 firms in 1998. The most restrictive sample requirement is that firms must have earnings forecast data available on I/B/E/S, which limits the sample to only those firms followed by analysts and by I/B/E/S, and therefore slants our sample toward larger firms. Requiring that two-year-ahead earnings forecasts be positive eliminates an average of only six firms per year among the set of firms covered by I/B/E/S. Collectively, our data requirements (particularly the requirement for ten consecutive years of data on Compustat) create the potential for survivorship bias.[6] Thus, our inferences may or may not generalize to younger firms, or to firms not followed by analysts.

Price Differentials from the Omission of Risk Premia

To test the pricing implications of residual income-based risk measures, we first measure the price differential for each sample firm-year as risk-free value (RFV_{it}) from applying the residual income valuation model using the risk-free rate of return as the discount rate, minus observed share price (P_{it}). Risk-free value is an estimate of firm value, assuming that investors are risk-neutral, and therefore, do not price risk.

We estimate RFV_{it} by modifying Equation (1) in two ways. First, in order to isolate the effects of risk from the estimation of risk-free value, we substitute the risk-free rate of return (r_f) for the required rate of return (r).[7] Thus, we use a risk-free rate to define abnormal earnings and discount them to present value. Second, to apply Equation (1) we must obtain expectations of earnings and book values over a finite forecast horizon and make assumptions about terminal value at the end of that horizon. To obtain earnings expectations we use I/B/E/S mean financial analysts' earnings per share forecasts for one and two years ahead, and we compute earnings expectations for years 3 through 5 by applying I/B/E/S analysts' consensus three- to five-year earnings per share growth rate estimates to their two-year-ahead forecasts. To avoid the unrealistic assumption of long-run negative earnings expectations, we eliminate firms with negative two-year-ahead earnings forecasts. To compute terminal value, we assume that the long-run nominal rate of growth in abnormal earnings (g) beyond year 5 equals 3 percent, the approximate long-run inflation rate. We compute terminal value as the present value of a growing perpetuity beginning with year 5 residual income. That is, we discount year 5 residual income using ($r_{ft} - g$), which is the contemporaneous, nominal, ten-year risk-free rate in period t (r_{ft}) less 3 percent. Our assumptions yield values of ($r_{ft} - g$) that are consistent with (but not equivalent to) the annual sample average estimates of ($r - g$) in Easton et al. (2000) for the years in which our samples overlap (1990–1998).[8] In addition, we estimate future book values of equity for each firm by projecting that the firm will maintain the dividend payout policy in year t (e.g., dividends paid as a percentage of earnings in year t) over the five-year forecast horizon.[9] We also assume that any additional share issues or repurchases over the horizon have neutral implications for abnormal earnings and that we can therefore safely omit them. Thus, we estimate RFV_{it} using the following equation:

---

[6] We increase the potential for survivorship bias and reduce our sample sizes in one set of multiple regressions in which we also require firms to report earnings in year $t + 2$, to control for a potential source of measurement error in analysts' two-year-ahead earnings forecasts.

[7] Our proxies for risk-free rates of return are the yields on ten-year U.S. government bonds, which we obtained from the Federal Reserve Bulletins (Table 1.35) as of April 1 of each sample year.

[8] The mean difference between our assumptions for ($r_{ft} - g$) and the Easton et al. (2000) estimates of ($r - g$) is less than 0.50 percent, although in three of our nine sample years the difference exceeds 1 percent in absolute value. Because r is greater than r_{ft} due to risk, we would expect Easton et al.'s (2000) ($r - g$) estimates to be greater than our ($r_{ft} - g$) assumptions, unless Easton et al.'s (2000) growth estimates are large (which they are). We explicitly control for different growth expectations in tests described later.

$$RFV_{it} = B_t + \sum_{i=1}^{4} \frac{E_t[X_{t+i} - r_{ft}B_{t+i-1}]}{(1 + r_{ft})^i} + \frac{E_t[X_{t+5} - r_{ft}B_{t+4}]}{(r_{ft} - g)(1 + r_{ft})^4} \tag{2}$$

We compute price differentials ($PDIFF_{it}$) as:

$$PDIFF_{it} = RFV_{it} - P_{it}, \tag{3}$$

where P_{it} is the price per share for firm i as of April 1 of each sample year for which we have analysts' earnings forecast data. Given that RFV omits the effect of risk on share price, PDIFF should be positive.

Table 1 provides annual sample averages of the inputs into the residual income valuation model (Panel A) and the resulting outputs from the valuation model (Panel B). Panel A indicates that average book values per share for our sample firms declined steadily from \$19.95 to \$14.18 per share over 1990 to 1998, despite the decline in average dividend payout rates from 58 percent to 29 percent during that time. We use I/B/E/S earnings forecasts over the five-year valuation horizon to compute expected future ROE as year t expected earnings divided by year t − 1 expected book value of equity. Averaging across the nine years of this study, these data indicate that expected ROE increases over the five-year forecast horizon, from an average of 12.2 percent for one-year-ahead forecasts to an average of 14.1 percent for five-year-ahead forecasts. The increase in average expected ROE over the five-year horizon suggests that analysts commonly expect future earnings to grow at a faster rate than book value of equity. In addition, analysts' expectations for future ROE over each of the five-year forecast horizons steadily trended upward over the sample period, from an average of roughly 11 percent in 1990 to roughly 17 percent in 1998. Concurrently, risk-free rates fell from 8.8 percent in 1990 to 5.6 percent in 1998. The six rightmost columns of Panel A report the mean forecasts of abnormal ROE, computed as forecasts of ROE minus the prevailing risk-free rates. These data show a dramatic increase in mean expected abnormal ROE over the study period, from 2.2 percent in 1990 to 11.3 percent in 1998.

As shown in Table 1, Panel B, risk-free values and share prices also increased from 1990 to 1998. This is not surprising, in light of the increasing expected abnormal ROE during this period. As expected, the sample average PDIFF is positive each year. In 1990 to 1992 when the average expected abnormal ROE is less than 4 percent per year, the average PDIFF is between \$6 and \$11. In 1996 to 1998, when the average abnormal ROE is 8 percent or more, the average PDIFF climbs to between \$22 and \$47.

The pricing differentials (PDIFF) reported in Table 1, Panel B, might seem large relative to observed prices, especially in later years. To calibrate the reasonableness of our PDIFF estimates, we compute the expected risk premia that the PDIFF estimates imply for each firm-year. PDIFF is isomorphic to the present value of the difference between the risk-free interest rate and the implicit risk-adjusted expected return for each firm (i.e., the risk premium implicit in expected returns). Therefore, we first estimated the implicit expected

---

[9] When year t earnings are negative, we compute the dividend payout rate as year t dividends divided by the mean expected earnings over the five-year earnings forecast horizon.

TABLE 1
Annual Means of Residual Income Valuation Model Inputs and Outputs

Panel A: Residual Income Valuation Model Inputs[a]

Year	Number of Firms	Book Value per Share	Dividend Payout Rate (%)	ROE Forecasts for Years (%):					Risk-Free Rates (%)	Implied AROE Forecasts for Years (%):					
				t + 1	*t +2*	*t + 3*	*t + 4*	*t + 5*		*t + 1*	*t + 2*	*t + 3*	*t + 4*	*t + 5*	*Mean*
1990	665	19.95	58	10.1	11.1	11.1	11.2	11.5	8.8	1.3	2.3	2.3	2.4	2.7	2.2
1991	773	18.46	56	9.0	10.5	10.9	11.1	11.7	8.0	1.0	2.5	2.9	3.1	3.7	2.6
1992	802	18.11	68	9.5	11.3	11.2	11.8	11.9	7.5	2.0	3.8	3.7	4.3	4.4	3.7
1993	822	16.45	52	11.1	12.5	12.6	13.2	13.3	6.0	5.1	6.5	6.6	7.2	7.3	6.6
1994	908	15.36	44	11.9	13.4	12.6	14.0	13.9	7.0	4.9	6.4	5.6	7.0	6.9	6.2
1995	898	15.50	38	13.3	14.3	13.7	15.0	15.2	7.0	6.2	7.2	6.6	7.9	8.1	7.2
1996	966	15.16	34	13.9	14.9	14.5	15.3	15.3	6.5	7.4	8.4	8.0	8.8	8.8	8.3
1997	1,060	14.85	33	14.9	15.8	15.2	16.0	16.1	6.9	8.0	8.9	8.3	9.1	9.2	8.7
1998	1,073	14.18	29	16.3	17.5	15.9	17.3	17.6	5.6	10.7	11.9	10.3	11.7	12.0	11.3
Mean	885	16.45	46	12.2	13.5	13.1	13.9	14.1	7.0	5.2	6.4	6.0	6.8	7.0	6.3

(*continued on next page*)

TABLE 1 (continued)

Panel B: Annual Means of Share Prices and Residual Income Valuation Model Outputs[b]

Year	*Number of Firms*	*Risk-Free Value per Share (RFV)*	*Price per Share (P)*	*Price Differential per Share (PDIFF)*	*Implied Risk Premium (%)*
1990	665	25.96	19.09	6.86	2.0
1991	773	26.31	19.97	6.34	1.7
1992	802	32.10	21.26	10.84	2.1
1993	822	51.76	23.23	28.53	3.3
1994	908	38.70	22.01	16.69	2.8
1995	898	45.36	23.37	21.99	3.3
1996	966	51.64	26.09	25.56	3.2
1997	1,060	49.63	27.10	22.54	3.1
1998	1,073	83.60	36.64	46.99	3.2
Overall Means	885	45.01	24.31	20.70	2.7

---

[a] We collect ***book value (of common equity per share)*** as of the end of year t−1, immediately preceding April of year t, from Compustat. ***Dividend payout rates*** equal the annual dividend per share divided by actual earnings per share (both from I/B/E/S). If actual earnings are negative, then we deflate dividends by average expected earnings per share for one year ahead through five years ahead (from I/B/E/S). ROEt + i (i = 1–5) is the return on common equity derived from I/B/E/S i-period-ahead forecasts and t + i − 1 book values. We estimate book values for periods beyond the current period as prior book value plus earnings times 1 minus the dividend payout ratio. We winsorize ROE data at 100 percent. ***Risk-free rates*** are yields on ten-year U.S. government bonds as of April of the indicated year, obtained from Federal Reserve Bulletins (Table 1.35). Implied forecasts of AROE (abnormal ROE) for year t + 1 through t + 5 equal forecasted ROE minus the prevailing risk-free rate.

[b] We use the residual income model, risk-free rates of return, and Equation (2) to compute ***risk-free value***. ***Price*** is the actual price per share as of April 1 of each year for which we have earnings forecasts. ***Price differential*** is risk-free value less actual share price. ***Implied risk premia*** denote the annual mean risk premia, computed by estimating the risky interest rate implied by security prices and then subtracting the risk-free rate. We obtain the interest rate implied by security prices by solving for the interest rate in Equation (2) that would cause RFV to equal PRICE.

return in share price by solving for the discount rate necessary to equate RFV to PRICE (i.e., to set PDIFF = 0). We then computed the expected risk premium implied by each PDIFF by subtracting the risk-free rate from the implicit risk-adjusted expected return estimate. The final column in Table 1, Panel B reports that the annual mean risk premium implied by each PDIFF ranges from 1.7 percent to 3.3 percent over our sample years, with a mean of 2.7 percent. Our mean implied risk premium is identical to the mean reported in Gebhardt et al. (2001) for the 1975–1995 period, and is slightly below the mean 3.4 percent reported by Claus and Thomas (2001) for our sample years. Claus and Thomas (2001) impose fewer data restrictions, and their sample includes smaller firms on average, so one would expect a slightly larger risk premium for the firms in their sample. The mean annual correlation between the estimates of implied risk premia and PDIFF is 0.88. Although the PDIFF and the implied risk premia estimates are not independent (both sets of estimates are based on the same models, the same analysts' forecasts, etc.), the implied risk premia estimates provide some calibration that the magnitudes of PDIFF are consistent with the magnitudes of risk premia observed in Gebhardt et al. (2001) and Claus and Thomas (2001).[10]

Two Residual Income-Based Risk Measures

Our first residual income-based risk measure is the systematic risk in abnormal ROE (i.e., abnormal ROE beta), which we estimate as the covariation between abnormal ROE and a sample-wide index of abnormal ROE, analogous to the accounting earnings beta in Beaver et al. (1970). For each firm-year observation, we regress abnormal ROE (denoted AROE) on an equally weighted sample average abnormal ROE (denoted AvgAROE) over the rolling ten-year window preceding the date on which we compute risk-free value:

$$\mathrm{AROE}_{it} = a_0 + \beta_{\mathrm{AROE}} \times \mathrm{AvgAROE}_t + e_{it}. \tag{4}$$

The coefficient estimate β_{AROE} is our proxy for systematic risk in abnormal ROE for each firm-year for which we have sufficient data. We imply subscripts i and t on each risk measure, but we drop them to streamline exposition.[11]

Our second residual income-based risk measure is the total risk in residual income, which we measure as the standard deviation of abnormal ROE (denoted σ_{AROE}) over the ten years prior to each risk-free value calculation.[12] We base our estimates of β_{AROE} and σ_{AROE} on only ten annual observations, and the limited time-series diminishes the power of our tests to detect their pricing implications. Based on our hypothesis, we expect PDIFF to be positively associated with β_{AROE} and σ_{AROE}.

Additional Risk Factor Measures

We use market model beta as the primary benchmark for our residual income-based risk measures. We compute market model beta (denoted β_{RET}) using a standard market model and an equally weighted market index, based on the 60 monthly returns preceding each risk-free value calculation. We expect PDIFF to be positively related to β_{RET} because it is the traditional measure of systematic equity risk.

---

[10] We repeated our empirical tests of our two residual income-based risk measures by examining their associations with the implied risk premia estimates rather than PDIFFs, and obtained very similar results.

[11] We winsorize AROE at 0.50 for each sample firm-year before constructing the sample-wide AvgAROE. Our results are not sensitive to alternative winsorization values up to 1.00.

[12] We do not deflate by the mean to obtain a coefficient of variation because AROE is already deflated by the book value of common equity. Controlling for measures of scale (market value of equity, earnings growth) does not change any of our conclusions about associations of σ_{AROE} with PDIFF.

We also use the other two factors in the Fama and French (1992) three-factor model (firm size and the book-to-market ratio) as benchmarks for our residual income risk measures. As in Fama and French (1992), we measure size as the natural log of market value of equity (denoted LNSIZE) as of the beginning of each year. We measure the book-to-market ratio as the natural log of the ratio of book value of equity divided by the market value of equity (denoted LNBM) as of the beginning of each year. If PDIFF measures the cost of risk, then we expect PDIFF to be negatively related to LNSIZE and positively related to LNBM, *ceteris paribus*, consistent with Fama and French (1992).

Controls for Potential Sources of Measurement Error in PDIFF

Each set of assumptions we use to compute RFV (analyst forecasts, terminal value growth rates, and dividend payout) creates potential measurement error. If measurement error occurs in RFV (e.g., analysts forecast earnings with a bias), and the same error occurs in PRICE (i.e., investors forecast earnings with the same bias), then the differencing process (RFV − PRICE = PDIFF) removes the error. For additional tests of the robustness of the pricing implications of our two residual income risk measures, we control for measurement error that might exist differentially in RFV and PRICE.

Analyst Forecast Errors

Prior research has detected an optimism bias in analysts' earnings forecasts, particularly for longer-term forecasts of annual earnings (e.g., McNichols and O'Brien 1997). Upwardly biased analyst forecasts increase RFV. If share prices reflect less optimism bias than analysts' earnings forecasts, then PDIFF is biased upward as well.[13] Our tests might be confounded if our risk factors are associated with the analyst forecast errors that introduce measurement error in PDIFF. To assess the effect of analyst forecast errors on our results, we control for the combined errors in one- and two-year-ahead forecasts for each firm-year (denoted AFE).[14]

Terminal Value Growth Assumption

Our assumption that residual income grows 3 percent per year after year 5 could introduce error into PDIFF if market prices are based on different growth assumptions. Cross-sectional differences in expected long-run growth are likely to be positively related to risk. Therefore, our tests of the association between PDIFF and residual income risk measures could be confounded if PDIFF and our risk measures both capture cross-sectional differences in expected earnings growth.[15] To control for this potential measurement error, we also include expected future earnings growth (denoted GROWTH), measured as I/B/E/S analysts' consensus forecasts of three- to five-year earnings growth.

---

[13] As expected by the way in which we construct RFV, ranked analyst forecast errors are positively correlated with RFVs—the median correlation is 0.275, which is very close to the median annual correlation of ranked analyst forecast error and PDIFF (deflated by price), 0.270. This result implies that security prices do not fully reflect the bias in analysts' forecasts.

[14] Our inferences are unaffected if we control for only one-year-ahead forecast errors for each year in the sample.

[15] Our computations of RFV capture cross-sectional differences in residual income growth ***through*** year 5, but they do not capture cross-sectional differences in expected growth rates in residual income *beyond* year 5 (which we assume to be a cross-sectional constant 3 percent). Insofar as differences in terminal-period residual income growth rates affect share prices, they will therefore affect PDIFF. For firms for which our assumption of 3 percent expected growth in residual income beyond year 5 is understated, our computation of PDIFF will be understated, and vice versa. Terminal-period expected growth in residual income is likely to covary positively with risk. If so, then measurement error in our assumption of 3 percent growth in terminal-year residual income will most likely bias our tests against our hypothesis of a positive association between PDIFF and residual income risk.

Dividend Payout Assumption

We assume that each firm's dividend payout rate remains constant, which could cause measurement error in PDIFF if market prices are based on expected changes in dividend payout rates. This assumption is likely to introduce less measurement error than the other two sources, given that most firms maintain stable dividend payout rates, so we do not explicitly control for expected changes in dividend payout.[16]

Descriptive Statistics

Table 2 contains descriptive statistics for our two measures of residual income risk, β_{AROE} and σ_{AROE}, as well as our primary benchmark risk measure, β_{RET}. Mean (median) β_{RET} is approximately 1 during the years 1990 to 1992, but then drops to less than 0.80 in 1993 and years after.[17] Median β_{AROE} exhibits a more gradual downward trend from 0.84 in 1990 to 0.60 in 1998. Mean β_{AROE} equals 1 by construction. The mean (median) σ_{AROE} gradually increases from 0.08 (0.06) in 1990 to 0.10 (0.08) in 1998. Of the three risk measures, β_{AROE} exhibits the greatest degree of in-sample dispersion. For example, in most sample years (except for 1998), the cross-sectional standard deviation of β_{AROE} is roughly four times greater than that of β_{RET}. Interestingly, Beaver et al. (1970, 667) make the same observation, even though their accounting beta is based on total (not abnormal) earnings and their sample period (1947–1965) precedes ours by at least 25 years.

Spearman rank correlations among β_{RET}, β_{AROE}, and σ_{AROE} (not tabulated) indicate that our residual income risk measures are related to but are not substitutes for market model beta. The correlations between σ_{AROE} and β_{RET} are statistically significantly positive (at $p < 0.05$) in six out of nine years, although the correlations are modest, ranging from 0.06 (1998) to 0.11 (1997). These low correlations suggest that β_{RET} and β_{AROE} are not close substitutes. The correlations between σ_{AROE} and β_{RET} are statistically significantly positive each year and average 0.35, suggesting the two risk measures capture similar but not identical factors. Not surprisingly, the correlations between σ_{AROE} and β_{AROE} are significant and positive each year. A curious result in these correlations (and in the analogous correlations reported in Beaver et al. [1970]) is that σ_{AROE} is uniformly more positively associated with β_{RET} than with β_{AROE}.

IV. TESTS AND RESULTS

We assess the share price implications of our three risk measures with a series of portfolio and regression tests. In this section we describe these tests and the results.

Portfolio Analysis

In H1 we predict that PDIFF increases as residual income risk increases, all else equal. As a preliminary analysis, we group our sample firms into ten portfolios each year based

---

[16] Supplemental analysis (not tabulated) suggests that assuming constant dividend payout introduces little error in PDIFF, in part because our sample selection criteria tilt the sample toward larger, more mature firms that likely have stable dividend policies. However, the sample does include firms with zero dividend payout policies. For example, in the 1990 sample of 665 firms, 82 (12.3 percent) had a 1990 payout ratio of zero. Four years later, 47 of the 54 firms common to both the 1990 and 1994 samples (87.1 percent) had maintained a zero payout ratio.

[17] Diagnostic analyses of our data reveal that this downward shift in β_{RET} is not attributable to changing sample composition over time or estimation procedures. In addition, supplemental analyses show a steady decline in market model betas over our sample period among all firms with 60 months of returns data in the intersection of the CRSP and I/B/E/S, as well as among all firms on CRSP with 60 months of returns data throughout our sample period.

TABLE 2
Annual and Pooled Descriptive Statistics for Three Alternative Measures of Risk: Systematic Risk in Stock Returns (β_{RET}), Systematic Risk in Abnormal ROE (β_{AROE}), and Standard Deviation in Abnormal ROE (σ_{AROE})[a]

Year	*Number of firms*	*Risk Measure*	*Mean*	*Standard Deviation*	*Minimum*	*Median*	*Maximum*
1990	665	β_{RET}	0.95	0.38	−0.25	0.96	4.11
		β_{AROE}	1.00	1.93	−9.06	0.84	11.31
		σ_{AROE}	0.08	0.06	0.02	0.06	0.38
1991	773	β_{RET}	0.99	0.42	−0.10	0.99	4.58
		β_{AROE}	1.00	2.21	−8.47	0.88	13.45
		σ_{AROE}	0.08	0.06	0.02	0.06	0.40
1992	802	β_{RET}	1.00	0.45	−2.18	0.95	4.54
		β_{AROE}	1.00	2.21	−8.73	0.83	14.39
		σ_{AROE}	0.09	0.07	0.00	0.06	0.43
1993	822	β_{RET}	0.78	0.50	−0.24	0.74	4.96
		β_{AROE}	1.00	2.57	−11.32	0.76	13.19
		σ_{AROE}	0.09	0.07	0.00	0.07	0.46
1994	908	β_{RET}	0.78	0.50	−1.22	0.74	3.61
		β_{AROE}	1.00	2.52	−8.36	0.79	14.05
		σ_{AROE}	0.09	0.07	0.00	0.07	0.48
1995	898	β_{RET}	0.78	0.56	−1.44	0.72	6.11
		β_{AROE}	1.00	2.39	−9.11	0.72	11.57
		β_{AROE}	0.10	0.07	0.01	0.07	0.45
1996	966	β_{RET}	0.72	0.61	−0.79	0.61	7.24
		β_{AROE}	1.00	2.86	−15.70	0.59	14.79
		σ_{AROE}	0.10	0.08	0.01	0.08	0.47
1997	1,060	β_{RET}	0.69	0.74	−5.29	0.57	9.81
		β_{AROE}	1.00	2.62	−7.86	0.58	14.59
		σ_{AROE}	0.11	0.08	0.01	0.08	0.48
1998	1,073	β_{RET}	0.71	1.96	−60.76	0.66	3.08
		β_{AROE}	1.00	2.54	−8.51	0.60	15.04
		σ_{AROE}	0.10	0.08	0.01	0.08	0.44
Pooled Across Years	7,967	β_{RET}	0.80	0.89	−60.76	0.78	9.81
		β_{AROE}	1.00	2.47	−15.70	0.73	15.04
		σ_{AROE}	0.10	0.07	0.00	0.07	0.48

---

[a] β_{RET} is a market model beta estimated using monthly stock returns and equally weighted market index returns over the 60 months prior to the year in which we compute price differentials from actual prices and Equation (2) (CRSP data). β_{AROE} is a proxy for systematic risk in abnormal ROE obtained by regressing abnormal ROE on the sample-wide average abnormal ROE over the ten-year period preceding the year in which we compute price differentials (Equation (4); Compustat data). σ_{AROE} is the standard deviation of abnormal ROE over a ten-year period preceding the year in which we compute price differentials (Compustat data). We winsorize abnormal ROEs at 50 percent prior to estimating β_{AROE} and σ_{AROE}.

on deciles of β_{AROE} and σ_{AROE}. To measure the magnitude of the average risk-related discount implicit in share price across deciles of β_{AROE} and σ_{AROE}, we computed mean PDIFFs for each decile for each year, 1990 to 1998. Grouping firms into deciles limits the power of statistical tests by collapsing hundreds of firms into ten portfolios each year. Moreover, these portfolio tests do not control explicitly for potentially confounding covariates (e.g., potentially correlated measurement error in PDIFF). Despite these limitations, we report the portfolio test results because the decile mean PDIFFs provide some insight into the likely magnitude (in dollars per share) of the effect of residual income risk on share prices. Prior studies report cross-sectional averages of implied rates of return or risk premia, but they do not provide dollar estimates of the economic effects of ignoring risk (e.g., Claus and Thomas 2001; Gebhardt et al. 2001).

Table 3 reports the median β_{AROE} and σ_{AROE} for each decile across all nine sample years and the across-year average PDIFF for each annual decile mean (in dollars per share).

TABLE 3

Mean Price Differentials (PDIFF) across Deciles of Systematic Risk in Residual Income (β_{AROE}) and Total Risk in Residual Income (σ_{AROE})[a]

Decile	*Median β_{AROE} across Years*	*Decile Mean PDIFF Averaged across Years*	*Median σ_{AROE} across Years*	*Decile Mean PDIFF Averaged across Years*
1 (lowest risk)	−3.00	18.64	0.02	14.05
2	−0.88	18.92	0.03	16.73
3	−0.12	16.98	0.04	18.20
4	0.23	18.07	0.05	20.46
5	0.59	17.50	0.06	21.63
6	0.95	20.48	0.07	22.57
7	1.38	22.03	0.10	22.15
8	1.98	21.87	0.12	22.76
9	2.95	25.14	0.16	26.81
10 (highest risk)	6.18	27.40	0.26	21.59
Annual means across deciles	—	20.70	—	20.70
Spearman correlation	—	0.79	—	0.81
Probability	—	<0.005	—	<0.005

---

[a] PDIFF denotes risk-free value (computed from the residual income model using risk-free rates of return; Equation (3) minus share price). σ_{AROE} is a proxy for systematic risk in abnormal ROE obtained by regressing abnormal ROE on the sample-wide average abnormal ROE over the ten-year period preceding each year in which we compute the residual income model (Equation (4)). σ_{AROE} is the standard deviation of abnormal ROE over a ten-year period preceding each residual income model computation. In estimating β_{AROE} and σ_{AROE}, we winsorize abnormal ROEs at 50 percent.

The results in Table 3, Columns 2 and 3 indicate that, averaging across years, the mean PDIFF is \$18.64 for the firms in the lowest β_{AROE} decile, and \$27.40 for firms in the highest β_{AROE} decile (nearly 47 percent larger). These magnitudes indicate that if the average stock in the lowest β_{AROE} decile has the sample mean share price of \$24.31 (see Table 1, Panel B), then it would be priced at \$42.95 in a hypothetical environment in which investors are risk-neutral and all else is held equal. Likewise, the average stock in the highest β_{AROE} decile would be priced at \$51.71. Consistent with our expectations that price differentials are increasing in β_{AROE}, the Spearman rank correlation between decile ranks of β_{AROE} and decile mean PDIFF is 0.79 ($p < 0.005$). In eight out of nine years (1990 is the exception), the Spearman rank correlations between decile ranks of β_{AROE} and mean PDIFF are positive and significant at $p < 0.10$ or better (not tabulated).[18]

The results in Table 3, Columns 4 and 5 show that the mean PDIFF across years for the lowest σ_{AROE} decile firms is \$14.05, whereas it is \$21.59 for the highest σ_{AROE} decile firms (more than 53 percent larger). Consistent with price differentials increasing in σ_{AROE}, the Spearman rank correlation between decile ranks of σ_{AROE} and decile mean PDIFF is 0.81 ($p < 0.005$). In addition, in every sample year, the Spearman rank correlation between decile ranks of σ_{AROE} and decile mean PDIFF is positive and significant at $p < 0.05$ or lower (not tabulated).

Regression Tests

We estimate a series of annual cross-sectional regressions to test whether PDIFF is positively associated with β_{AROE} and σ_{AROE}. We first estimate univariate regressions of PDIFF on the ranks (within each sample year) of each residual income risk measure. We also estimate univariate regressions of PDIFF on ranked β_{RET} to benchmark the explanatory power of β_{AROE} and σ_{AROE}. The univariate regressions take the form:

$$PDIFF_{it}/P_{it} = a + cRiskMeasureRank_{it} + e_{it}, \tag{5}$$

where $RiskMeasureRank_{it}$ denotes the annual within-sample rank of the risk factor regressor ($Rank\beta_{AROE}$, $Rank\sigma_{AROE}$, or $Rank\beta_{RET}$), for firm i in year t. We use the within-sample rank of each risk measure (rather than the estimated value of each risk measure) to mitigate the effects of measurement error in the risk measures. We scale $PDIFF_{it}$ by P_{it} to control for possible effects of per-share scale on our regression results.[19]

We then estimate multiple regressions including all three ranked risk measures to assess their incremental implications for PDIFF:

$$PDIFF_{it}/P_{it} = a + c_1 Rank\beta_{AROEit} + c_2 Rank\sigma_{AROEit} + d_1 Rank\beta_{RETit} + e_{it}, \tag{6}$$

where all variables are as defined in Equation (5).

---

[18] Note that the median β_{AROE} is -3.0 among the 1st decile firms, but the mean PDIFF for these firms is greater than the mean PDIFF for the 3rd, 4th, and 5th deciles. Subsequent analyses (not tabulated) reveal that the 1st decile firms have higher-than-average σ_{AROE} and β_{RET}. These firms—with abnormal ROEs that seem to hedge market movements but stock returns that mirror market movements—are an interesting group in their own right. Another alternative explanation is that these firms have β_{AROE} measured with great error. We also formed the deciles based on PDIFF, similar to the technique in Beaver et al. (1980) and the reverse regression approach in Beaver et al. (1987), and found that the correlation between PDIFF and β_{AROE} was only slightly higher.

[19] Our primary motivation for deflating PDIFF by price in the univariate regressions is to obtain results that will be comparable with the multiple regression results that include control variables that are correlated with scale. When we do not deflate PDIFF by price, the univariate and multiple regression results for our two residual income risk measures are very similar to those reported in the paper. In addition, untabulated tests using unranked winsorized data yield similar conclusions.

Finally, in our most complete multiple regressions, we regress PDIFF on three classes of variables: (1) the two residual income risk measures (our primary variables of interest); (2) the three Fama and French (1992) risk factors (β_{RET}, LNSIZE, and LNBM); and (3) two controls for potential measurement error in PDIFF (GROWTH and AFE):

$$\begin{aligned} PDIFF_{it}/P_{it} &= a + c_1 Rank\beta_{AROEit} + c_2 Rank\sigma_{AROEit} + d_1 Rank\beta_{RETit} \\ &\quad + d_2 LNSIZE_{it} + d_3 LNBM_{it} + f_1 GROWTH_{it} + f_2 AFE_{it} + e_{it}, \qquad (7) \end{aligned}$$

where all variables are as defined previously.

Hypothesis 1 predicts that the coefficient estimates on each ranked residual income risk measure will be positive in univariate regressions. In the multiple regressions, the c_1 and c_2 coefficient estimates on $Rank\beta_{AROE}$ and $Rank\sigma_{AROE}$ in Equation (6) will indicate whether the capital markets price these risk measures incrementally to systematic risk in returns ($Rank\beta_{RET}$). The c_1 and c_2 coefficient estimates from Equation (7) will indicate whether the capital markets price $Rank\beta_{AROE}$ and $Rank\sigma_{AROE}$ incrementally to the three factors in the Fama and French (1992) model, after controlling for analyst forecast errors and growth.

Table 4 contains the results of the regression estimations. For each set of annual cross-sectional regressions, we report the results for the three univariate regressions (in the first three rows for each year); the multiple regressions with $Rank\beta_{AROE}$, $Rank\sigma_{AROE}$, and $Rank\beta_{RET}$ (in the fourth rows); and the multiple regressions with the three Fama and French (1992) factors and the controls (in the fifth rows). At the end of the table, we aggregate the results over time by reporting a Z-statistic for each intertemporal mean t-statistic.[20]

The first row of results for each year in Table 4 reports that the univariate regression coefficient estimates for $Rank\beta_{AROE}$ are positive and significant, as hypothesized, in six of nine years and aggregated across years ($Z = 4.42$). However, these annual regressions' adjusted R^2s range from -0.1 percent to 1.7 percent, indicating that $Rank\beta_{AROE}$ explains little of the variation in PDIFF. Overall, these results suggest that systematic risk in residual income has a positive, but modest, association with price differentials.

The second row of results for each year in Table 4 shows that the coefficients on $Rank\sigma_{AROE}$ in the univariate regressions are positive and significant, as hypothesized, in eight of nine years and aggregated across years ($Z = 7.46$). The adjusted R^2 on $Rank\sigma_{AROE}$ ranges from 0.2 percent to 7.2 percent, and is the highest of all three risk measures in five of nine years. These results suggest that total volatility in residual income is positively associated with price differentials.

As reported in the third row of results for each year in Table 4, the univariate regression coefficient estimates on $Rank\beta_{RET}$ are positive and significant (at $p < 0.05$ or lower) in seven of nine sample years and aggregated over time ($Z = 4.49$). The adjusted R^2 ranges from -0.1 percent to 9.2 percent in the annual regressions. These results are consistent with the traditional notion that the capital markets price systematic risk in returns.

The fourth row in each set of annual results in Table 4 reports the coefficient estimates from multiple regressions in which all three ranked risk measures compete to explain PDIFF. The coefficient estimates on $Rank\beta_{AROE}$ are significantly positive in only two of nine years; the pooled Z-statistic is 2.71. The coefficients on $Rank\sigma_{AROE}$ are consistently positive and significant (at $p < 0.05$ or lower) in all but one of the nine sample years; the

---

[20] We calculate this Z-statistic as $Z = \bar{t}\sqrt{(N - 1)}/stddev(t)$, where $\bar{t}$ is the average t-statistic over all sample years, N is the number of sample years (9), and stddev(t) is the standard deviation of our sample t-statistics (Barth 1994; Collins et al. 1995.) We do not report results from regressions pooled over time because they suffer from cross-sectionally dependent residuals and overstated t-statistics (Bernard 1987).

TABLE 4
Annual Cross-Sectional Regressions of Percent Pricing Differentials (PDIFF) on Residual Income Risk Measures, Fama and French Risk Factors, and Control Variables[a]

Table entries are estimated coefficients, t-statistics (in parentheses), and adjusted R^2 (last column).

Year	(number of firms)	Intercept	Ranked Abnormal ROE Beta ($Rank\beta_{AROE}$)	Ranked Abnormal ROE Volatility ($Rank\sigma_{AROE}$)	Ranked Market Return Beta ($Rank\beta_{RET}$)	Log of Market Value of Equity (LNSIZE)	Log of the Book-to-Market Ratio (LNBM)	I/B/E/S Long-Run Growth Estimates (GROWTH)	Combined One- and Two-Year Ahead Analyst Forecast Errors (AFE)	Adjusted R^2 (%)
1990	(665)	0.34 (5.18)	−0.01 (−0.57)							−0.1
	(665)	0.22 (3.31)		0.03 (1.59)						0.2
	(665)	0.28 (4.16)			0.01 (0.60)					−0.1
	(665)	0.27 (3.00)	−0.02 (−0.95)	0.03 (1.63)	−0.00 (−0.09)					0.1
	(643)	−0.08 (−0.41)	−0.00 (−0.86)	−0.00 (−0.35)	−0.00 (−0.33)	0.00 (0.06)	−0.22 (−4.81)	1.60 (2.07)	0.00 (5.13)	10.9
1991	(773)	0.21 (4.02)	0.02 (1.48)							0.2
	(773)	0.08 (1.61)		0.05 (4.33)						2.3
	(773)	0.24 (4.57)			0.01 (0.84)					−0.0
	(773)	0.09 (1.23)	0.01 (0.63)	0.05 (4.13)	−0.01 (−0.83)					2.1
	(753)	0.06 (0.44)	0.00 (0.95)	0.00 (1.49)	−0.00 (−0.89)	−0.05 (−2.71)	−2.39 (−6.48)	0.89 (1.64)	0.00 (6.88)	16.8
1992	(802)	0.33 (5.72)	0.03 (2.66)							0.8
	(802)	0.22 (3.82)		0.06 (4.96)						2.9
	(802)	0.37 (6.46)			0.02 (1.79)					0.3
	(802)	0.16 (2.06)	0.02 (1.55)	0.06 (4.10)	−0.01 (−0.07)					2.9
	(782)	0.17 (1.06)	0.00 (1.06)	0.00 (1.70)	0.00 (0.14)	−0.05 (−2.51)	−0.26 (−5.98)	0.24 (0.42)	0.00 (6.98)	16.3
1993	(822)	1.10 (15.04)	0.03 (2.00)							0.4
	(822)	0.76 (10.65)		0.12 (7.72)						6.7
	(822)	0.81 (11.35)			0.10 (6.76)					5.2
	(822)	0.56 (5.86)	0.01 (0.34)	0.09 (5.36)	0.07 (4.34)					8.5
	(799)	0.33 (1.50)	0.00 (0.32)	0.00 (2.59)	0.00 (3.72)	−0.05 (−1.88)	−0.24 (−4.41)	2.20 (3.13)	0.00 (6.01)	18.2

(continued on next page)

TABLE 4 (continued)

Year (number of firms)	Intercept	Ranked Abnormal ROE Beta ($Rank\beta_{AROE}$)	Ranked Abnormal ROE Volatility ($Rank\sigma_{AROE}$)	Ranked Market Return Beta ($Rank\beta_{RET}$)	Log of Market Value of Equity (LNSIZE)	Log of the Book-to-Market Ratio (LNBM)	I/B/E/S Long-Run Growth Estimates (GROWTH)	Combined One- and Two-Year Ahead Analyst Forecast Errors (AFE)	Adjusted R^2 (%)
1994 (908)	0.73 (14.53)	0.01 (1.44)							0.1
(908)	0.53 (10.85)		0.06 (6.03)						3.8
(908)	0.51 (10.33)			0.06 (7.72)					4.7
(908)	0.37 (5.40)	0.01 (0.65)	0.04 (3.74)	0.05 (4.99)					6.1
(867)	−0.01 (−0.07)	0.00 (0.56)	0.00 (1.24)	0.00 (3.23)	−0.01 (−0.68)	−0.05 (−1.10)	2.07 (4.31)	0.00 (7.66)	15.5
1995 (898)	0.83 (14.93)	0.03 (2.94)							0.9
(898)	0.58 (10.80)		0.09 (8.38)						7.2
(898)	0.53 (9.97)			0.10 (9.57)					9.2
(898)	0.32 (4.49)	0.01 (1.22)	0.06 (5.03)	0.08 (7.10)					12.0
(851)	−0.05 (−0.32)	0.00 (1.86)	0.00 (4.86)	0.00 (4.17)	−0.01 (−0.36)	0.12 (2.52)	1.47 (2.86)	0.00 (7.78)	18.6
1996 (966)	0.87 (15.78)	0.04 (4.26)							1.7
(966)	0.72 (13.35)		0.07 (7.49)						5.4
(966)	0.67 (12.53)			0.08 (8.64)					7.1
(966)	0.42 (5.92)	0.03 (2.73)	0.04 (4.20)	0.07 (6.37)					9.8
(899)	0.57 (3.31)	0.00 (3.21)	0.00 (3.43)	0.00 (0.76)	−0.06 (−3.39)	0.16 (3.30)	3.43 (8.11)	0.00 (4.05)	19.6
1997 (1,060)	0.75 (11.66)	0.05 (4.34)							1.7
(1,060)	0.50 (8.07)		0.09 (8.99)						7.0
(1,060)	0.54 (8.62)			0.08 (8.25)					6.0
(1,060)	0.23 (2.87)	0.02 (2.15)	0.06 (5.57)	0.06 (5.22)					9.6
(968)	0.98 (4.95)	0.00 (3.65)	0.00 (3.32)	−0.00 (−0.13)	−0.15 (−6.94)	−0.03 (−0.42)	2.42 (4.70)	0.00 (5.24)	19.5
1998 (1,073)	1.28 (18.42)	0.03 (2.84)							0.5
(1,073)	0.95 (14.10)		0.09 (8.54)						6.3
(1,073)	0.92 (13.75)			0.10 (9.03)					7.0
(1,073)	0.68 (7.66)	0.01 (0.91)	0.06 (4.98)	0.07 (6.21)					9.4
(1,034)	1.10 (6.25)	0.00 (2.36)	0.00 (2.66)	−0.00 (−2.34)	−0.12 (−6.29)	0.10 (1.91)	7.11 (15.15)	0.00 (3.16)	32.8

TABLE 4 (continued)

Year (number of firms)	Intercept	Ranked Abnormal ROE Beta ($Rank\beta_{AROE}$)	Ranked Abnormal ROE Volatility ($Rank\sigma_{AROE}$)	Ranked Market Return Beta ($Rank\beta_{RET}$)	Log of Market Value of Equity (LNSIZE)	Log of the Book-to-Market Ratio (LNBM)	I/B/E/S Long-Run Growth Estimates (GROWTH)	Combined One- and Two-Year Ahead Analyst Forecast Errors (AFE)	Adjusted R^2 (%)
Z-statistics computed from the sample of nine annual t-statistics		(4.42)							
			(7.46)						
				(4.49)					
		(2.71)	(10.26)	(3.33)					
		(2.85)	(4.36)	(1.16)	(−3.12)	(−1.28)	(2.97)	(10.32)	

---

[a] Table 4 reports coefficient estimates, t-statistics, and adjusted R²s from the following annual cross-sectional regressions:

Univariate regressions (rows one to three each year):

$$PDIFF_{it}/P_{it} = a + cRiskMeasureRank_{it} + e_{it} \qquad (5)$$

Multiple regressions without controls (row four each year):

$$PDIFF_{it}/P_{it} = a + c_1Rank\beta_{AROEit} + c_2Rank\sigma_{AROEit} + d_1Rank\beta_{RETit} + e_{it} \qquad (6)$$

Multiple regressions with controls (row five each year):

$$PDIFF_{it}/P_{it} = a + c_1Rank\beta_{AROEit} + c_2Rank\sigma_{AROEit} + d_1Rank\beta_{RETit} + d_2LNSIZE_{it} + d_3LNBM_{it} + f_1GROWTH_{it} + f_2AFE_{it} + e_{it} \qquad (7)$$

PDIFF denotes risk-free value (computed from the residual income model using risk-free rates of return; Equation (3)) minus share price. P denotes share price of firm i as of April 1 for each year t for which we have earnings forecasts. β_{AROE} is a proxy for systematic risk in abnormal ROE obtained by regressing abnormal ROE on the sample-wide average abnormal ROE over the ten-year period preceding each year in which we compute risk-free value from the residual income model (Equation (4)). σ_{AROE} is the standard deviation of abnormal ROE over a ten-year period preceding each residual income model computation. In estimating β_{AROE} and σ_{AROE}, we winsorized abnormal ROEs at 50 percent. β_{RET} is a market model beta estimated using monthly stock returns and equally weighted market index returns over the 60 months prior to the year in which we compute risk-free value from the residual income model (CRSP data). We transform all three risk measures into ranks based on their within-sample rank each year. LNSIZE is a control variable measured as the natural log of firm size. LNBM is the natural log of the book-to-market ratio. GROWTH controls for expected future earnings growth, measured as I/B/E/S consensus analysts' forecasts of three- to five-year earnings growth. AFE controls for analyst forecast errors, measured as the combined forecast errors for one- and two-year-ahead forecasts for each year in the sample.

pooled Z-statistic is 10.26. The coefficient estimates on Rankβ_{RET} are positive and significant (at $p < 0.05$ or better) in six of nine sample years; the pooled Z-statistic is 3.33. These results indicate that total volatility in residual income has the most robust incremental relation with price differentials, even beyond that of systematic risk in returns. These results also suggest that systematic risk in residual income does not have robust incremental relevance for share prices beyond either systematic risk in returns or total volatility in residual income.

We report the results for the multiple regressions including all three Fama and French (1992) factors and the control variables in the fifth row in each set of annual results in Table 4. In these regressions, the coefficient estimates on Rankβ_{AROE} are significantly positive in only three of nine years (pooled Z = 2.85). The coefficients on Rankσ_{AROE} diminish in significance, but remain significantly positive in six out of nine years (pooled Z = 4.36). These results confirm that total volatility in residual income has robust incremental association with price differentials, even beyond the Fama and French (1992) risk factors, and after controlling for potential measurement error attributable to AFE and GROWTH. The results for Rankβ_{RET} are the most sensitive to the inclusion of the other risk factors and the controls. The coefficient estimates on Rankβ_{RET} are significantly positive in only three of nine sample years, and are not significant in the aggregate (Z = 1.16). This latter result is not surprising—others (e.g., Fama and French 1992) have shown that after controlling for firm size and the book-to-market ratio, beta has little explanatory power for returns. As expected, the coefficients on LNSIZE are significant and negative in six of the nine sample years (pooled Z = –3.12). The coefficients on LNBM are unstable—negative and significant from 1990 to 1993, and then positive and significant in 1995, 1996, and 1998 (pooled Z = –1.28). The coefficients on the control variables GROWTH and AFE are positive and significant in most sample years, indicating that long-run expected earnings growth and analyst forecast errors provide incremental explanatory power for price differentials beyond our two measures of residual income risk and the three Fama and French (1992) risk factors.[21]

Supplemental Analysis—Risk Measures Based on Reported Net Income

We next investigate whether risk measures based on reported net income exhibit associations with price differentials that are similar to those found for risk measures based on residual income. We are interested in this question because it is easier to measure and use reported net income than residual income.

Our residual income risk measures are highly positively correlated with analogous risk measures based on reported net income. Thus, when we repeat the regression analyses reported in Table 4, substituting β_{ROE} for β_{AROE} (i.e., systematic risk in income measured using ROE rather than abnormal ROE) and substituting σ_{ROE} for σ_{AROE} (i.e., total volatility in ROE rather than abnormal ROE), we find that both provide comparable explanatory power for price differentials (results not tabulated).[22] Comparable to the multiple regression results in Table 4, we find that the capital market prices total volatility in ROE incrementally to systematic risk in returns, and that systematic risk based on ROE does not provide consistent incremental explanatory power for share prices. Although residual income-based risk measures are consistent with theory (Feltham and Ohlson 1999), as a practical matter, risk measures based on reported income are simpler to compute than risk measures based on residual income, and both have similar share price implications.

---

[21] The significant associations between PDIFF and GROWTH and AFE imply that future research may be warranted to (1) develop measures of RFV that correct for both long-run earnings growth and analysts' forecast errors, and (2) examine the extent to which share prices fully reflect long-run earnings growth and analysts' forecast errors.

[22] Beaver et al. (1970) use reported earnings deflated by market prices in constructing total and systematic accounting risk measures. We deflate by an accrual accounting measure, book value of equity.

V. SUMMARY AND CONCLUSIONS

Risk plays a fundamental but not yet well-understood role in residual income valuation. Feltham and Ohlson (1999) emphasize the role of risk in residual income valuation and point out that the capital markets should price nondiversifiable (systematic) variability inherent in expected future residual income. However, it is not clear exactly how one should incorporate risk into empirical tests or practical applications of the residual income valuation framework. Consequently, empirical researchers have used different methods of incorporating risk in empirical applications of residual income valuation, with different results.

The objective of this study is to enhance our understanding of the role of risk in residual income-based valuation. We develop a new accounting-based measure of the effect of risk on share price, using the difference between observed share price and risk-free value, measured using the residual income model and risk-free rates of return. This new measure—which we denote the *price differential*—captures the magnitude of the discount for risk implicit in share prices. We also develop two accounting-based measures of risk in residual income—systematic risk and total volatility. Our tests examine whether our risk measures explain the cross-section of price differentials, and whether the capital markets price these two risk measures incrementally to the three Fama and French (1992) risk factors (i.e., beta, firm size, and book-to-market ratios).

Our results suggest that the capital markets price systematic risk in residual income. Abnormal ROE beta is significantly positively related to price differentials in univariate regressions, but provides only limited explanatory power, indicating that abnormal ROE beta is, at best, a weak indicator of risk. Multiple regressions confirm these results—we find that in only three of nine years does the capital market price abnormal ROE beta incrementally to the three Fama and French (1992) risk factors and two control variables for measurement error. These tests have low power because of the inherent difficulty in estimating abnormal earnings betas, so it may be fruitful to explore more powerful ways to estimate abnormal (or raw) earnings betas.

Our results suggest that total volatility in residual income is robustly positively associated with price differentials, and that the capital markets price total volatility in residual income incrementally to the three Fama and French (1992) risk factors. In fact, our findings suggest that total volatility in residual income has more explanatory power for price differentials than systematic risk in residual income, beta, firm size, or the book-to-market ratio. Despite the fact that we do not measure total volatility in residual income using explicit systematic risk factors, total volatility in residual income nevertheless captures priced risk factors. These results help explain why corporate managers interested in maximizing share value might prefer smooth income series over income volatility.

If our accounting- and market-based risk proxies are valid measures of risk, then our findings suggest that our measure of the discount for risk implicit in share price, which is based on the residual income model and prevailing risk-free rates, is a theoretically defensible, computationally simple, and empirically valid measure of the impact of risk on share prices. Our findings also suggest that one can assess firm risk using volatility and covariation in abnormal earnings, consistent with the residual income model's focus on abnormal earnings as the fundamental valuation attribute. In future research, we plan to test whether accounting risk proxies (especially abnormal ROE volatility) explain share prices in initial public offering settings in which market data are not available, and thus, in which one cannot use the market model or the Fama and French (1992) model to assess and price firm risk.

REFERENCES

Barth, M. 1994. Fair value accounting: Evidence from investment securities and the market valuation of banks. *The Accounting Review* 69 (January): 1–25.

Beaver, W., P. Kettler, and M. Scholes. 1970. The association between market-determined and accounting-determined measures of risk. *The Accounting Review* 45 (October): 654–682.

———, R. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting and Economics* 2 (March): 3–28.

———, ———, and S. Ryan. 1987. The information content of security prices: A second look. *Journal of Accounting and Economics* 9 (July): 139–157.

Bernard, V. L. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* 25 (Spring): 1–48.

———. 1994. Accounting-based valuation methods, determinants of market-to-book ratios, and implications for financial statement analysis. Working paper, University of Michigan, Ann Arbor, MI.

———. 1995. The Feltham-Ohlson framework: Implications for empiricists. *Contemporary Accounting Research* 11 (Spring): 733–747.

Botosan, C., and M. Plumlee. 2001. Estimating expected cost of equity capital: A theory-based approach. Working paper, University of Utah, Salt Lake City.

———, and ———. 2002. A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research* 40 (March): 21–40.

Claus, J., and J. Thomas. 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. *Journal of Finance* 56 (October): 1629–1666.

Collins, J., D. Shackelford, and J. Wahlen. 1995. Bank differences in the coordination of regulatory capital, earnings and taxes. *Journal of Accounting Research* 33 (Autumn): 263–291.

Dechow, P., A. Hutton, and R. Sloan. 1999. An empirical assessment of the residual income valuation model. *Journal of Accounting and Economics* 26 (January): 1–34.

Easton, P., G. Taylor, P. Shroff, and T. Sougiannis. 2000. Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. *Journal of Accounting Research* 40 (June): 657–676.

Edwards, E. O., and P. W. Bell. 1961. *The Theory and Measurement of Business Income*. Berkeley, CA: University of California Press.

Fama, E., and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (June): 427–465.

Feltham, G. A., and J. A. Ohlson. 1995. Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research* 11 (Spring): 689–731.

———, and ———. 1999. Residual earnings valuation with risk and stochastic interest rates. *The Accounting Review* 74 (April): 165–183.

Francis, J., P. Olsson, and D. Oswald. 2000. Comparing the accuracy and explainability of dividend, free cash flow and abnormal earnings equity valuation estimates. *Journal of Accounting Research* 38 (Spring): 45–70.

———, ———, and ———. 2001. Using mechanical earnings and residual income forecasts in equity valuation. Working paper, Duke University.

Gebhardt, W., C. M. C. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39 (June): 135–176.

McNichols, M., and P. C. O'Brien. 1997. Self-selection and analyst coverage. *Journal of Accounting Research* 35 (Supplement): 167–199.

Ohlson, J. A. 1995. Earnings, book values and dividends in equity valuation. *Contemporary Accounting Research* 11 (Spring): 661–687.

Peasnell, K. V. 1982. Some formal connections between economic values and yields and accounting numbers. *Journal of Business Finance and Accounting* 9 (October): 361–381.

Preinreich, G. A. D. 1938. Annual survey of economic theory: The theory of depreciation. *Econometrica* 6: 219–241.

THE ACCOUNTING REVIEW
Vol. 78, No. 1
January 2003
pp. 353–371

Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing

Patricia M. Fairfield
Georgetown University
J. Scott Whisenant
University of Houston
Teri Lombardi Yohn
Georgetown University

ABSTRACT: Prior research reveals that the accrual component of profitability is less persistent than the cash flow component, and that investors fail to fully appreciate their differing implications for future profitability (Sloan 1996). However, accruals are a component of growth in net operating assets as well as a component of profitability. Just as we can disaggregate profitability into accruals and cash flows from operations, we can disaggregate growth in net operating assets into accruals and growth in long-term net operating assets. We find that, after controlling for current profitability, both components of growth in net operating assets—accruals and growth in long-term net operating assets—have equivalent negative associations with one-year-ahead return on assets. This result is consistent with conservative accounting and diminishing marginal returns on investments. We also find that, after controlling for current profitability, the market appears to equivalently overvalue accruals and growth in long-term net operating assets relative to their association with

---

We appreciate the comments of Bill Baber, Mary Barth, Sudipta Basu, Ken Cavalluzzo, Dennis Chambers, Agnes Cheng, Dan Collins, Bala Dharan, Paul Hribar, Prem Jain, Son-Hyon Kang, Charles Lee, Maureen McNichols, Michael Mosebach, James Ohlson, Stephen Penman, Srinivasan Sankaraguruswamy, Richard Sloan, Tom Stober, Hong Xie, and Stephen Zeff. This study has also benefited from the comments of workshop participants at Emory University, The George Washington University, Hong Kong University of Science and Technology, New York University, Rice University, Stanford University, Texas Christian University, University of Arkansas, University of Houston, University of Maryland, and the 2001 American Accounting Association Annual Meeting. This research was supported in part by a Steers Faculty Research Fellowship awarded by the McDonough School of Business.

Submitted November 2000
Accepted July 2002

one-year-ahead ROA. Our evidence suggests that the accrual anomaly documented in Sloan (1996) is a special case of what could be viewed as a more general growth anomaly.

Keywords: *accrued earnings; growth; earnings quality; financial statement analysis; market mispricing.*

Data Availability: *Data are available from sources identified in the paper.*

I. INTRODUCTION

Research on financial statement analysis documents that financial statement components and ratios provide information about future profitability (Ou and Penman 1989; Ou 1990; Abarbanell and Bushee 1997; Fairfield and Yohn 2001). Research on market efficiency concludes that investors misprice such fundamental historical financial information as firm size (Fama and French 1992), recent earnings surprises (Freeman and Tse 1989; Bernard and Thomas 1989), and past sales growth (Lakonishok et al. 1994). Evidence of differential persistence and apparent market mispricing of the accrual and cash flow components of profitability contributes to both streams of research (Sloan 1996; Collins and Hribar 2000; DeFond and Park 2001; Xie 2001). Sloan (1996) finds that the accrual component of current profitability is less persistent than the cash flow component. He also reports that investors fail to fully price the differing implications of the accrual and cash flow components of current profitability for one-year-ahead profitability. In particular, investors appear to overvalue accruals.

Ohlson (1995) and Feltham and Ohlson (1995), however, show that future profitability and firm value depend on growth in net operating assets as well as current profitability. We can disaggregate growth in net operating assets into two components: accruals and growth in long-term net operating assets—just as we can disaggregate profitability into accruals and cash flows from operations. Prior research on the persistence and valuation of cash flows vs. accruals focuses on the role of accruals as a component of profitability, but overlooks the role of accruals as a component of growth in net operating assets. Our study probes the extent to which the differential persistence and valuation of accruals documented in prior research can be explained by the role of accruals as a component of growth in net operating assets. In other words, we investigate whether the differential persistence and apparent mispricing of accruals that Sloan (1996) documents applies more broadly to growth in net operating assets.

Recognizing that accruals are a component of growth in net operating assets, as well as a component of profitability, is important in interpreting prior evidence that accruals are less persistent than cash flows and apparently overvalued by investors. We expect that *given current levels of profitability*, defined as current-period return on assets (ROA), growth in net operating assets will depress one-year-ahead ROA. This is because both conservative accounting principles and diminishing marginal returns to increased investment tend to reduce profitability for growing firms. We expect this negative incremental effect of growth in net operating assets on one-year-ahead ROA to result in a negative correlation between accruals (a component of growth in net operating assets) and one-year-ahead ROA, conditional on current ROA. We also expect the conditional relation between accruals and one-year-ahead ROA to be equivalent to the conditional relation between growth in long-term net operating assets and one-year-ahead ROA.

Consistent with our predictions, we show that one-year-ahead ROA is negatively associated with both components of growth in net operating assets—accruals and growth in long-term net operating assets—after controlling for current ROA. Furthermore, conditional on current ROA, there is no statistically significant difference in the negative relations between the two components of growth in net operating assets and one-year-ahead ROA. We interpret these results as suggesting that the lower persistence of accruals relative to cash flows documented in Sloan (1996) and Xie (2001) is a specific manifestation of a more general negative incremental relation between one-year-ahead ROA and growth in net operating assets.[1] To the extent that the lower persistence of accruals arises because accruals are a component of growth in net operating assets, the lower persistence of accruals relative to cash flows from operations is more likely to result from the conservative bias in accounting principles or the lower rate of economic profits that result from diminishing marginal returns to new investment opportunities, or both. At the same time, the lower persistence of accruals is less likely to result from other features of accruals, such as their susceptibility to manipulation by management.

If the overvaluation of accruals relative to cash flows documented in Sloan (1996) is attributable to the market's misunderstanding of the incremental effect on one-year-ahead ROA of growth in net operating assets, then we would expect to find a similar overvaluation of the second component of growth in net operating assets—growth in *long-term* net operating assets. Sloan (1996, 305) concludes that earnings expectations reflected in stock prices deviate from rational expectations as if investors naively fixate on aggregate earnings, and fail to appreciate the differential persistence of different components of earnings. Investors may similarly misinterpret the implications for future profitability of growth in long-term net operating assets because this growth has a weaker association with future profitability than does current aggregate earnings. Moreover, as we show later, after controlling for current ROA, accruals and growth in long-term net operating assets have similar implications for one-year-ahead ROA. To the extent that investors misprice accruals because they fail to understand the implications of accruals for one-year-ahead ROA, investors may similarly misprice the growth in long-term net operating assets if they fail to understand its similar implications for one-year-ahead ROA.

We investigate whether the apparent market inefficiency with respect to investors' overvaluation of accruals documented in Sloan (1996) appears to be a special case of investors' difficulty in valuing growth in net operating assets. We use the Mishkin (1983) test for market mispricing and find that, after controlling for the valuation implications of current ROA, investors appear to overweight the valuation implications of both accruals and growth in long-term net operating assets for one-year-ahead ROA. In addition, we find no statistically significant difference in the severity of mispricing of the two forms of growth. Our evidence suggests that the market's apparent mispricing of accruals relates to investors' inability to assess correctly the implications of growth in net operating assets for future profitability, whether the growth is in current or long-term net operating assets.

The next section of the paper develops the hypotheses on the implications of accruals and growth for one-year-ahead ROA and market mispricing. Section III describes the variables used in the analysis. Section IV provides the empirical evidence on the implications of accruals and growth for one-year-ahead ROA and market mispricing. Section V presents the conclusions.

II. HYPOTHESIS DEVELOPMENT

Sloan (1996) regresses future profitability, defined as one-year-ahead ROA, on the accrual and cash flow components of current ROA. He concludes that accrued earnings are less persistent than cash earnings. However, just as profitability can be disaggregated into accruals and cash flows, growth in net operating assets can be disaggregated into accruals and growth in long-term net operating assets. Consequently, it is not clear whether the lower persistence of accruals documented in Sloan (1996) is attributable to its role as a component of profitability or to its role as a component of growth in net operating assets. To address this question, we investigate whether, for firms with similar levels of current profitability (i.e., similar ROA_t), those firms that invest more heavily in operating assets during year t have lower ROA in year t + 1 relative to similar firms that invest less heavily in operating assets during year t.

---

[1] Xie (2001, 365) finds that the persistence of cash flows from operations significantly exceeds the persistence of normal accruals, which in turn exceeds the persistence of abnormal accruals.

We expect both diminishing marginal returns on investment (Stigler 1963, 54) and conservative accounting (Penman 2001, 565) to contribute to this negative incremental relation between future profitability and growth in net operating assets. Diminishing marginal returns on investments arise when firms exploit their most profitable investment opportunities before undertaking less profitable investments. Alternatively, increasing marginal returns on divestment arise when firms divest their least profitable investments.[2] Penman (2001, 561) shows that a conservative bias in accounting procedures results in investments that appear relatively less profitable in early years and more profitable in later years, biasing accounting rates of return on new investments downward relative to the returns on existing investments. Thus, both diminishing marginal returns and conservative accounting suggest that firms investing more in net operating assets during year t will experience lower one-year-ahead ROA relative to other firms, *ceteris paribus* (i.e., for firms with similar ROA).

Absent theory suggesting that the effects of diminishing marginal returns on investments or conservative accounting should differ between current and long-term net operating asset growth, we do not expect accruals and growth in long-term net operating assets to differ in their implications for one-year-ahead ROA. Our first hypothesis is:

> **H1:** Accruals and growth in long-term net operating assets have equivalent incremental negative relations with one-year-ahead ROA, after controlling for current ROA.

Prior research presents evidence that the market appears to overvalue accruals relative to its implications for one-year-ahead ROA (Sloan 1996; Xie 2001; DeFond and Park 2001). If the overvaluation of accruals is attributable to the market's misunderstanding of the incremental effects of growth in net operating assets on one-year-ahead ROA, then we would expect a similar overvaluation of the second component of growth in net operating assets—growth in *long-term* net operating assets. Thus, our second hypothesis is:

> **H2:** After controlling for current ROA, mispricing of growth in long-term net operating assets is equivalent to the mispricing of accruals.

III. RESEARCH DESIGN

Sample Selection

Our sample consists of firms with required financial statement and returns data for the 30-year period from 1964 to 1993. This sample period ensures replication of Sloan's (1996) results using a similar research design and 30-year sample period.[3] We extract the financial data from the annual database of Standard & Poor's 2000 Compustat Industrial and Merged Industrial Research files. We obtain the stock price and share data from the Center for Research in Security Prices (CRSP) 2001 monthly stock price database.

---

[2] For ease of exposition, we refer to both scenarios as diminishing marginal returns to investment.

[3] Sloan (1996) uses financial data for the 30-year period from 1962 to 1991. However, we eliminate 1962 and 1963 from our study period because data requirements reduce the number of observations in those years to unreasonable sample sizes (less than 20 firms) on which to run annual cross-sectional regressions. Therefore, we add two years, 1992 and 1993, to yield a 30-year sample period as in Sloan (1996).

We cannot directly estimate the accrual and cash flow components of operating income for years prior to the enactment of SFAS No. 95 in 1988. Thus, we follow Sloan (1996) and use an indirect (balance sheet) approach that relies on the articulation between changes in working capital and accruals.[4]

We eliminate the sample firm-years without sufficient financial data to compute operating income, accruals, cash flows, or net operating assets. There are 38,865 firm-years with stock price data and sufficient data to estimate the financial statement variables. Since we use the indirect method to estimate accruals, we also eliminate firm-years that would most likely induce noise into the measurement of accruals: (1) Compustat footnote codes suggest that working capital components are estimated since the company reports an unclassified balance sheet (2,563 firm-years); (2) Compustat footnote codes suggest that managers of the firm made voluntary accounting changes that affect either operating income or working capital accounts (2,635 firm-years); and (3) goodwill (Compustat item 240) increased from year to year since, in purchase business combinations, changes in the balance sheet accounts may not articulate with income statement accruals (706 firm-years). The sample selection process results in a final sample of 32,961 firm-years with available stock price and financial statement data in prior, current, and subsequent years.

Measurement of Variables

The explanatory variables in the regression analyses include current ROA, accruals, cash flows, growth in net operating assets, and growth in long-term net operating assets, and we define each in turn.

Consistent with Sloan (1996), we define ROA as operating income deflated by contemporaneous average total assets (for parsimony, we suppress cross-sectional firm subscripts throughout the study):

$$ROA_t = \frac{OPINC_t}{AVG(TA_{t-1} + TA_t)},$$

where:

OPINC = operating income after depreciation and amortization (Compustat item 178), and

TA = total assets (Compustat item 6).

---

[4] Studies document nonarticulation problems with the indirect approach to estimating the accrual component of earnings (e.g., Drtina and Largay 1985; Bahnson et al. 1996; Hribar and Collins 2002). For example, Hribar and Collins (2002) document that merger and acquisition activities, foreign currency translations, and divestitures introduce significant measurement error into accruals estimated using the indirect method. Our inferences are unaffected if we eliminate the approximately 12,000 firm-years affected by these transactions, however. We report results without deleting these firm-years to enhance comparability with Sloan (1996) and to provide larger sample sizes for our year-by-year regressions.

Using the indirect method, we define accruals (ACC) as growth (net change) in operating working capital accounts (other than tax liabilities) minus current-period depreciation and amortization expense.[5] That is, we define accruals as:[6]

$$ACC_t = GrWC_t - DEPAMORT_t,$$

where:

$$GrWC_t = (\Delta AR_t + \Delta INV_t + \Delta OTHERCA_t) - (\Delta AP_t + \Delta OTHERCL_t);$$

and
ΔAR = change in accounts receivable (Compustat item 2);
ΔINV = change in inventories (Compustat item 3);
$\Delta OTHERCA$ = change in other current assets (Compustat item 68);
ΔAP = change in accounts payable (Compustat item 70);
$\Delta OTHERCL$ = change in other current liabilities (Compustat item 72); and
$DEPAMORT$ = depreciation and amortization expense (Compustat item 14).

We calculate *cash flows from operations* (CFO) by subtracting ACC from operating income (OPINC) as follows:

$$CFO_t = OPINC_t - ACC_t.$$

Following Sloan (1996), we deflate ACC and CFO by contemporaneous average total assets.

We define *growth in net operating assets* (GrNOA) as annual change in net operating assets:

$$GrNOA_t = NOA_t - NOA_{t-1},$$

where *net operating assets* (NOA) is operating assets (excluding cash) minus operating liabilities:[7]

$$NOA_t = AR_t + INV_t + OTHERCA_t + PPE_t + INTANG_t + OTHERLTA_t - AP_t - OTHERCL_t - OTHERLTL_t$$

---

[5] We remove taxes payable and deferred taxes because operating income excludes tax expense. The income measure also excludes interest expense, suggesting that we should exclude interest payable as well. Interest accruals arising from differences between periodic interest payments and recorded interest expense are recorded in the net book value of debt, and thus we exclude them from the estimate of accruals. However, Compustat reports interest accruals arising from timing differences at fiscal year-end as either "debt in current liabilities" or "current liabilities—other." When Compustat reports interest accruals in "debt in current liabilities," they are appropriately excluded from accruals. When Compustat reports interest accruals in "current liabilities—other," they are incorrectly included in accruals. Our results are robust to the exclusion of "other current liabilities" from the estimate of accruals.

[6] This method excludes accruals attributable to income-statement elements below operating income. Other studies using SFAS No. 95 data (e.g., Collins and Hribar 2002) define accruals as income before extraordinary items and discontinued operations, minus cash flows from operations taken from the statement of cash flows. Fairfield et al. (1996) find that income statement items lower on the income statement have less predictive ability for future return on equity. Barth et al. (2001) find that the (non-operating) accruals excluded under the indirect method have less ability to predict future cash flows from operations. Thus, our results apply only to operating accruals. Non-operating accruals may be less persistent than both operating accruals and cash flows from operations.

[7] We focus on growth in net *operating* assets to be consistent with our definition of ROA that is based on operating income and our definition of accruals based on growth in operating working capital.

where:

AR = accounts receivable (Compustat item 2);
INV = inventories (Compustat item 3);
OTHERCA = other current assets (Compustat item 68);
PPE = net property, plant, and equipment (Compustat item 8);
INTANG = intangibles (Compustat item 33);
OTHERLTA = other long-term assets (Compustat item 69);
AP = accounts payable (Compustat item 70);
OTHERCL = other current liabilities (Compustat item 72); and
OTHERLTL = other long-term liabilities (Compustat item 75).

Our final explanatory variable—*growth in long-term net operating assets* (GrLTNOA)—is defined as growth in net operating assets other than accruals:

$$\text{GrLTNOA}_t = \text{GrNOA}_t - \text{ACC}_t.$$[8]

We deflate GrNOA and GrLTNOA by contemporaneous average total assets. Because GrLTNOA equals GrNOA minus ACC, GrLTNOA reduces to changes in operating asset and liability elements excluded from ACC (i.e., ΔPPE and ΔINTANG exclusive of depreciation and amortization expense, ΔOTHERLTA, and ΔOTHERLTL).

IV. RESULTS

Descriptive Statistics

Table 1 provides descriptive statistics on one-year-ahead ROA, current ROA, accruals, cash flows, growth in net operating assets, and the components of growth in net operating assets (growth in working capital, depreciation and amortization expense, and growth in long-term net operating assets). Panel A shows that the mean current ROA for the sample firms is 11.4 percent. As documented in prior studies (Sloan 1996), mean accruals are negative (−0.02), suggesting that accruals, on average, decrease income.

Panel C of Table 1 shows that this is because the (negative) effect of depreciation and amortization expense outweighs the (positive) growth in working capital. Panel C also shows that long-term net operating assets are growing more than three times faster than working capital. The finding highlights that failing to consider the effects of growth in not only current, but also long-term net operating assets may lead to incorrect inferences about why there is a differential relation between accruals and cash flows for future profitability and mispricing. If the results hold for not only accruals, but also growth in long-term net operating assets, then the differential relations are less likely to relate to earnings management and more likely to relate to the effect of growth in net operating assets in general.

Table 2 reports correlations among one-year-ahead ROA, current ROA, accruals, cash flows, growth in total net operating assets, growth in working capital, depreciation and amortization expense, and growth in long-term net operating assets. As expected, one-year-ahead ROA is more positively correlated with cash flows from operations than with accruals. This evidence is consistent with accruals reflecting both income (which is associated with higher future ROA) and growth in net operating assets (which as we show later has a negative incremental association with future ROA).

---

[8] Our results are robust if we remove depreciation and amortization expense from ACC and include it with GrLTNOA.

TABLE 1
Descriptive Statistics on One-Year Ahead ROA, Current ROA, and Current Growth in Net Operating Assets and the Components of Current ROA and Current Growth in Net Operating Assets

Variable[a]	*Mean*	*Standard Deviation*	*Median*	*First Quartile*	*Third Quartile*
Panel A: One-Year-Ahead and Current ROA					
ROA_{t+1}	0.111	0.110	0.107	0.063	0.162
ROA_t	0.114	0.112	0.109	0.065	0.165
Panel B: Components of Current ROA					
ACC_t	−0.020	0.101	−0.027	−0.065	0.017
CFO_t	0.134	0.132	0.135	0.079	0.195
Panel C: Total and Components of Growth in Net Operating Assets, where GrNOA = GrWC − DEPAMORT + GrLTNOA					
$GrNOA_t$	0.070	0.158	0.055	−0.003	0.126
$GrWC_t$	0.024	0.095	0.014	−0.013	0.055
$DEPAMORT_t$	−0.044	0.029	−0.038	−0.054	−0.027
$GrLTNOA_t$	0.090	0.117	0.068	0.033	0.122

---

[a] Number of observations = 32,961 firm-years from 1964 to 1993. The variables are defined as:

ROA_{t+1} = return on assets, defined as operating income after depreciation and amortization at time t+1 divided by average total assets at time t+1;
ROA_t = return on assets, defined as operating income after depreciation and amortization at time t divided by average total assets at time t;
ACC_t = accruals, defined as the change in current operating assets minus the change in current operating liabilities (exclusive of tax liabilities) minus depreciation and amortization expense, divided by average total assets;
CFO_t = cash flows from operations, defined as ROA minus ACC;
$GrNOA_t$ = the one-year growth in net operating assets, where net operating assets equals operating assets minus operating liabilities, divided by average total assets;
$GrWC_t$ = the one-year growth in noncash working capital, where working capital is defined as current operating assets minus current operating liabilities, divided by average total assets;
$DEPAMORT_t$ = current period depreciation and amortization expense (shown as a negative value), divided by average total assets; and
$GrLTNOA_t$ = one-year growth in long-term net operating assets, divided by average total assets.

Table 2 shows positive unconditional correlations among current-period growth in net operating assets, current period ROA, and one-year ahead ROA. The positive unconditional correlation between current-period growth and one-year ahead ROA is not inconsistent with our hypothesis that growth in net operating assets will have a negative *incremental* effect on one-year ahead ROA, because H1 is conditional on controlling for the level of current profitability. As the correlations in Table 2 show, firms with higher levels of profitability in the current period tend to both: (1) invest more in net operating assets, and (2) have higher levels of profitability one year ahead. Hypothesis 1 posits that, *after controlling for current profitability*, the growth in net operating assets during year t is negatively associated

TABLE 2

Correlations among ROA and Growth in Net Operating Assets and Their Components[a]
(Spearman coefficients in the upper triangle; Pearson coefficients in the lower triangle)

Panel A: Correlations among One-Year-Ahead ROA, Current ROA, and Current Growth in Net Operating Assets[a]

Variables	ROA_{t+1}	ROA_t	ACC_t	CFO_t	$GrNOA_t$	$GrWC_t$	$DEP\ AMORT_t$	$GrLTNOA_t$
ROA_{t+1}	1.00	0.79*	0.18*	0.54*	0.20*	0.19*	0.01	0.13*
ROA_t	0.78*	1.00	0.28*	0.62*	0.34*	0.30*	−0.03*	0.23*
ACC_t	0.14*	0.24*	1.00	−0.47*	0.64*	0.92*	−0.42*	0.03*
CFO_t	0.55*	0.66*	−0.57*	1.00	−0.19*	−0.41*	0.26*	0.16*
$GrNOA_t$	0.13*	0.26*	0.67*	−0.30*	1.00	0.71*	−0.03*	0.71*
$GrWC_t$	0.13*	0.23*	0.96*	−0.54*	0.71*	1.00	−0.11*	0.17*
$DEPAMORT_t$	−0.04*	−0.07*	−0.36*	0.21*	−0.02*	−0.07*	1.00	0.35*
$GrLTNOA_t$	0.06*	0.14*	0.04*	0.09*	0.77*	0.13*	0.28*	1.00

Panel B: Correlations among GrLTNOA and Components of GrWC[b]

Variables	$GrLTNOA_t$	ΔAR_t	ΔINV_t	$\Delta OTHER\ CA_t$	ΔAP_t	$\Delta OTHER\ CL_t$
$GrLTNOA_t$	1.00	0.24*	0.24*	0.14*	0.25*	0.20*
ΔAR_t	0.23*	1.00	0.34*	0.12*	0.45*	0.34*
ΔINV_t	0.19*	0.36*	1.00	0.12*	0.42*	0.26*
$\Delta OTHERCA_t$	0.04*	0.03*	0.00	1.00	0.06*	0.20*
ΔAP_t	0.24*	0.48*	0.45*	0.13*	1.00	0.19*
$OTHERCL_t$	0.12*	0.21*	0.15*	0.20*	0.09*	1.00

---

* Denotes correlations are significantly different from 0, p-value < 0.01.

[a] Number of observations = 32,961 firm-years from 1964 to 1993.

The variables in Panel A are defined in Table 1.

[b] $GrWC_t = (\Delta AR_t + \Delta INV_t + \Delta OTHERCA_t) - (\Delta AP_t + \Delta OTHERCL_t)$ where:

ΔAR = one-year change in accounts receivable (Compustat item 2);

ΔINV = one-year change in inventories (Compustat item 3);

$\Delta OTHERCA$ = one-year change in other current assets (Compustat item 68);

ΔAP = one-year change in accounts payable (Compustat item 70); and

$\Delta OTHERCL$ = one-year change in other current liabilities (Compustat item 72).

All variables are deflated by contemporaneous average total assets.

with firms' profitability in year t+1. Consistent with H1, untabulated analysis reveals that the unconditional correlation between current period growth in net operating assets and the *change* in one-year-ahead ROA is −0.30.

In Panel B of Table 2, we report the correlations between growth in long-term net operating assets and the components of growth in working capital. Growth in long-term net operating assets is positively correlated with growth in each working capital component

(ΔAR, 0.24; ΔINV, 0.24; ΔOTHERCA, 0.14; ΔAP, 0.25; ΔOTHERCL, 0.20), suggesting that current operating assets and liabilities grow in support of growth in long-term net operating assets.

The Incremental Effect of Growth in Net Operating Assets on One-Year-Ahead ROA (H1)

In this section, we investigate the incremental effects of the two components of growth in net operating assets on one-year-ahead ROA, after controlling for current profitability. As in Fama and MacBeth (1973), we estimate annual regressions and report t-statistics based on the means and standard deviations of the parameter estimates obtained in 30 annual regressions.[9]

The first column in Table 3 provides evidence on the association between one-year-ahead ROA and growth in net operating assets (GrNOA) after controlling for current ROA:

$$ROA_{t+1} = \alpha_0 + \alpha_1 GrNOA_t + \alpha_2 ROA_t + \mu_{t+1}. \quad (1)$$

As expected, we find a significantly negative coefficient on GrNOA, after controlling for current ROA. Panel C of Table 3 reveals that the coefficient on GrNOA is significantly negative ($p < 0.01$) in 20 of the 30 annual regressions, and significantly positive in only one regression. Thus, in contrast to the positive unconditional univariate correlation between growth in net operating assets and one-year-ahead ROA (in Table 2), Table 3 shows that growth in net operating assets has the expected negative incremental association with one-year-ahead ROA, after controlling for current ROA. In other words, for firms with similar profitability in year t, firms investing relatively more in net operating assets experience lower profitability in year t+1. This result is consistent with accounting conservatism and diminishing marginal returns depressing the one-year-ahead ROA on the new investment.

In H1 we hypothesize that, after controlling for current ROA, the two components of growth in net operating assets (accruals and growth in long-term net operating assets) will have statistically equivalent negative relations with one-year-ahead ROA. To test H1, we disaggregate GrNOA into accruals and growth in long-term net operating assets, allowing each of the components to have different coefficients. If, after controlling for current profitability, both components of balance sheet growth have similar implications for one-year-ahead ROA, then the coefficient on GrLTNOA (β_1) will not differ significantly from the coefficient on ACC (β_2) in the following model:

$$ROA_{t+1} = \beta_0 + \beta_1 GrLTNOA_t + \beta_2 ACC_t + \beta_3 ROA_t + e_{t+1}. \quad (2)$$

The results of estimating Equation (2) appear in Table 3. We find the expected negative coefficients on both growth in long-term net operating assets and accruals.[10] Panel B shows the results are consistent with H1. After controlling for current ROA, the coefficient on accruals, β_2, does not differ significantly from the coefficient on growth in long-term net

---

[9] To ensure consistency with Sloan (1996), we also regress one-year-ahead ROA on current accruals and cash flows from operations. The mean estimated coefficient on cash flows from operations (0.74) significantly exceeds the mean coefficient on accruals (0.68) at $p < 0.01$. Thus, consistent with Sloan (1996), we find that the accrual component of ROA is less persistent than the cash flow component when we do not control for growth in net operating assets.

[10] The *total effect of one dollar of accruals* on future ROA is positive, because accruals are a component of ROA as well as a component of growth in net operating assets (0.73 = 0.79 − 0.06), whereas the total effect of one dollar of growth in long-term net operating assets is negative (−0.04).

TABLE 3
Regressions of One-Year-Ahead ROA on Growth in Net Operating Assets, after Controlling for Current ROA (tests of H1)

$$ROA_{t+1} = \alpha_0 + \alpha_1 GrNOA_t + \alpha_2 ROA_t + u_{t+1} \quad (1)$$

$$ROA_{t+1} = \beta_0 + \beta_1 GrLTNOA_t + \beta_2 ACC_t + \beta_3 ROA_t + e_{t+1} \quad (2)$$

Panel A: Estimation Results for Each Equation[a]

	Equation (1)		*Equation (2)*	
Variable	*Estimated Coefficient*	*t-statistic*	*Estimated Coefficient*	*t-statistic*
Intercept	0.02	(11.79)***	0.02	(10.75)***
$GrNOA_t$	−0.05	(−7.09)***	—	—
$GrLTNOA_t$	—	—	−0.04	(−5.25)***
ACC_t	—	—	−0.06	(−5.61)***
ROA_t	0.78	(54.63)***	0.79	(54.42)***
Average adjusted R^2	0.62		0.62	

Panel B: Test of Differences in Coefficients on GrLTNOA and ACC ($\beta_1 = \beta_2$ in Equation (2))

$GrLTNOA_t = ACC_t$	t-statistic (p-value)	1.65 (0.11)

Panel C: Parameter Tests by Year

	Equation (1)		*Equation (2)*	
	No. Years	*Implied by Result*	*No. Years*	*Implied by Result*
Tests based on	20	GrNOA < 0	7	ACC < GrLTNOA
significance level of	9	GrNOA = 0	20	ACC = GrLTNOA
p-value < 0.01	1	GrNOA > 0	3	ACC > GrLTNOA

***,**, and * Denotes significance at the 1, 5, and 10 percent levels (two-tailed test), respectively.

[a] The coefficients and adjusted R^2 are averages obtained from the 30 annual regressions from 1964 to 1993 (32,961 firm-years). The t-statistics are based on the averages and standard deviations of the 30 parameter estimates obtained in the annual regressions.

The variables are defined in Table 1.

operating assets, β_1 ($p = 0.11$)[11] Panel C reports the results of parameter tests by year. Using a 1 percent significance level, we find that the coefficient on GrLTNOA does not differ significantly from the coefficient on ACC in 20 of the 30 annual regressions. Of the remaining ten annual regressions, the coefficient on ACC is significantly larger than the coefficient on GrLTNOA in three years, and significantly smaller in seven years. The results reported in Table 3 also reveal that there is no increase in explanatory power from disaggregating the growth in net operating assets into its GrLTNOA and ACC components.

---

[11] In supplemental tests, we perform pooled analyses as well as analyses by industry groups, by ROA deciles, and after excluding extreme values identified using the Belsley et al. (1980) procedure. Our inferences are robust to these sensitivity tests. While the coefficients on growth in long-term net operating assets, accruals, and ROA differ across (two-digit SIC) industries and ROA deciles, the main result that accruals and growth in long-term net operating assets do not differ in their implications for one-year-ahead ROA, after controlling for current ROA, holds across the industries and ROA deciles, and after excluding extreme values.

The results from Equation (2) also imply that, after controlling for accruals as a component of growth in net operating assets, there is no significant difference between the persistence of accruals and cash flows. That is, the coefficients in the following model:

$$ROA_{t+1} = \delta_0 + \delta_1 GrNOA_t + \delta_2 ACC_t + \delta_3 CFO_t + v_{t+1}, \tag{2a}$$

may be inferred from the coefficients in Equation (2). Specifically, the coefficients in Table 3 imply that the coefficient on accruals, δ_2, does not differ significantly from the coefficient on cash flows from operations, δ_3.[12] This evidence that there is no significant difference between the persistence of accruals *vis-à-vis* cash flows (after controlling for growth in net operating assets) is not consistent with conjectures in prior research that the lower persistence of accruals is attributable to pervasive earnings management (Sloan 1996).

Market Pricing Tests (H2)

Our second hypothesis predicts that, if the overvaluation of accruals relative to cash flows documented in Sloan (1996) is attributable to the market's misunderstanding of the incremental effects of growth in net operating assets on one-year-ahead ROA, then the market's overvaluation of accruals will generalize to growth in long-term net operating assets. In Table 4, we provide evidence on the mispricing of accruals and growth in long-term net operating assets.

Mishkin (1983) develops a framework to test whether investors appear to rationally price publicly available information.[13] Our application of the Mishkin framework provides a statistical comparison between (1) a measure of the market's pricing of growth in long-term net operating assets and accruals (i.e., the market's *valuation coefficient* on growth in long-term net operating assets and accruals), and (2) a measure of the ability of growth in long-term net operating assets and accruals to predict one-year-ahead ROA (i.e., the *forecast coefficient* on growth in long-term net operating assets and accruals). That is, we compare investors' apparent pricing of growth in long-term net operating assets and accruals (i.e., valuation coefficients) to the implications of growth in long-term net operating assets and accruals for one-year-ahead ROA (i.e., forecasting coefficients). If the valuation coefficients significantly exceed the forecasting coefficients, then the Mishkin test suggests that investors overprice the implications of that component of growth for one-year-ahead ROA.

---

[12] Since ACC is a component of both ROA and GrNOA and is included as a separate explanatory variable in Equation (2), we can infer the coefficients on GrNOA, ACC, and CFO in Equation (2a) from the estimated coefficients in Equation (2). In Equation (2a), the coefficient on CFO equals the coefficient on ROA in Equation (2). That is, $\delta_3 = \beta_3 = 0.79$. In addition, the coefficient on GrNOA in Equation (2a) equals the coefficient on GrLTNOA in Equation (2). That is, $\delta_1 = \beta_1 = -0.04$. Finally, the coefficient on ACC in Equation (2a) equals the coefficient on ROA plus the coefficient on ACC and minus the coefficient on GrLTNOA in Equation (2). That is, $\delta_2 = \beta_3 + \beta_2 - \beta_1 = 0.79 + (-0.06) - (-0.04) = 0.77$. Since the coefficient on ACC (β_2) does not significantly differ from the coefficient on GrLTNOA (δ_1) in Equation (2), the coefficient on ACC (δ_2) does not significantly differ from the coefficient on CFO (δ_3) in Equation (2a).

[13] The Mishkin framework does not require a complete specification of the relation between variables in the forecasting equation and subsequent earnings (Mishkin 1983; Abel and Mishkin 1983). However, the framework does assume that security prices are efficient with respect to variables that are omitted from the equations and are correlated with the predictor variables. If one-year-ahead returns include a delayed response to information in correlated omitted variables, the weighting of the predictor variables in the valuation equations will include both the market's delayed adjustment to the correlated omitted variable(s) and the market's weighting of the predictor variables.

TABLE 4
Nonlinear Generalized Least Squares Estimation (the Mishkin Test) of the Market Pricing of Growth in Net Operating Assets and ROA with Respect to Their Implications for One-Year-Ahead ROA
(tests of H2)

Forecasting Equation:

$$ROA_{t+1} = \beta_0 + \beta_1 GrLTNOA_t + \beta_2 ACC_t + \beta_3 ROA_t + e_{t+1} \qquad (2)$$

Valuation Equation:[a]

$$BHAR_{t+1} = \alpha + \theta(ROA_{t+1} - \beta_0 - \beta_1^* GrLTNOA_t - \beta_2^* ACC_t - \beta_3^* ROA_t) + \varepsilon_{t+1} \qquad (3)$$

Panel A: Market Pricing of Growth in Net Operating Assets and ROA with Respect to Their Implications for One-Year-Ahead ROA

Forecasting Coefficients			*Valuation Coefficients*		
Parameter (variable)	*Estimate*	*Asymptotic Std. Error*	*Parameter (variable)*	*Estimate*	*Asymptotic Std. Error*
β_1(GrLTNOA)	−0.04	0.003	β_1^*(GrLTNOA)	0.03	0.009
β_2(ACC)	−0.06	0.004	β_2^*(ACC)	0.05	0.011
β_3(ROA)	0.79	0.004	β_3^*(ROA)	0.73	0.011

Panel B: Tests of Rational Pricing of Current GrLTNOA, ACC, and ROA

Null Hypotheses	*Likelihood Ratio Statistic*[b]	*Marginal Significance Level*
GrLTNOA, ACC, ROA: $\beta_1^* = \beta_1$ and $\beta_2^* = \beta_2$ and $\beta_3^* = \beta_3$	143.41	<0.001
GrLTNOA : $\beta_1^* = \beta_1$	14.55	<0.001
ACC : $\beta_2^* = \beta_2$	75.01	<0.001
ROA : $\beta_3^* = \beta_3$	67.41	<0.001
GrLTNOA, ACC: $\beta_1^* = \beta_2^*$ and $\beta_1 = \beta_2$	1.72	0.423

---

[a] Equations (2) and (3) are jointly estimated using an iterative generalized nonlinear least squares estimation procedure based on 32,961 firm-years from 1964 to 1993.

[b] The likelihood ratio statistic for the full model [2n log(SSRc/SSRu)] = 2 × 32,961 × Ln(13,007.848/12,979.580) = 143.41

All variables are defined in Table 1.

Specifically, we estimate a system in which the forecasting equation is Equation (2) and the valuation equation is as follows:

$$BHAR_{t+1} = \alpha + \theta(ROA_{t+1} - \beta_0 - \beta_1^* GrLTNOA_t - \beta_2^* ACC_t - \beta_3^* ROA_t) + \varepsilon_{t+1} \qquad (3)$$

where $BHAR_{t+1}$ is the size-adjusted abnormal stock return, computed as the difference between a firm's annual buy-and-hold return and the annual buy-and-hold return for the same 12-month period on the market-capitalization-based portfolio decile to which the firm belongs.

Equation (2) estimates the forecasting coefficients (βs) on current growth in long-term net operating assets, accruals, and ROA for predicting one-year-ahead ROA. Equation (3) estimates the valuation coefficients (β^*s) that the market appears to assign to growth in long-term net operating assets, accruals, and ROA. As Mishkin (1983) proposes, we estimate Equations (2) and (3) jointly, using an iterative generalized nonlinear least squares estimation procedure and proceeding in two stages.

In the first stage, we impose no constraints on βs and β^*s and jointly estimate Equations (2) and (3). To test whether the valuation coefficients (β^*s) are significantly different from the counterpart forecasting coefficients (βs) obtained in the first stage, we jointly estimate Equations (2) and (3) in the second stage after imposing the rational pricing constraints, $\beta_q^* = \beta_q$, where q = 1, 2, or 3. Mishkin (1983) shows that the following likelihood ratio statistic is distributed asymptotically as $\chi^2(q)$ under the null hypothesis, which, in our application, is that the market rationally prices all the explanatory variables in Equation (2) with respect to their implications for one-year-ahead ROA:

$$2n \log (SSR^c/SSR^u),$$

where:

q = number of rational pricing constraints;
n = the number of observations;
log = natural logarithm operator;
SSR^c = the sum of squared residuals from the constrained weighted system; and
SSR^u = the sum of squared residuals from the unconstrained weighted system.

Panel A of Table 4 reports the estimated coefficients for Equations (2) and (3) obtained from the first stage (unconstrained) estimation. The forecasting coefficients are those reported for Equation (2) in Table 3. Panel B reports tests of rational pricing. A test of the overall model results in a significant likelihood ratio statistic of 143.41 ($p < 0.001$), suggesting that we reject the null hypothesis that the market rationally prices the components of GrNOA and current ROA.

The valuation coefficient on GrLTNOA ($\beta_1^* = 0.03$) exceeds its forecasting coefficient ($\beta_1 = -0.04$), and is of the opposite sign. This suggests that the market appears to overvalue growth in long-term net operating assets relative to its ability to predict one-year-ahead ROA. To test whether this mispricing is statistically significant, we jointly estimate Equations (2) and (3) in the second stage, after imposing the rational pricing constraint (i.e., $\beta_1^* = \beta_1$). As reported in Panel B, the apparent overpricing of growth in long-term net operating assets is significant ($p < 0.001$) after controlling for the valuation implications of profitability.

We also reject the rational pricing of accruals. The valuation coefficient on accruals $\beta_2^* = 0.05$) is greater than its forecasting coefficient ($\beta_2 = -0.06$), suggesting that the market overvalues the implications of accruals for one-year-ahead ROA. Panel B of Table 4 reports that the likelihood ratio statistic rejects the null hypothesis that the market rationally prices accruals ($p < 0.001$) after controlling for the valuation implications of current profitability.

The valuation coefficient on ROA ($\beta_3^* = 0.73$) is smaller than the forecasting coefficient ($\beta_3 = 0.79$) suggesting that the market underprices ROA relative to its implications for one-year-ahead ROA, after controlling for growth in net operating assets. The likelihood ratio statistic of 67.41 reported in Panel B of Table 4 is significant ($p < 0.001$), indicating that we can reject the null hypothesis of rational pricing.[14]

We also test whether the extent of the apparent overpricing differs between growth in long-term net operating assets and accruals. From the results in Table 3 we know that there is no significant difference between the coefficients on accruals and long-term growth in net operating assets in the forecasting equations. To test for differential severity in mispricing, we estimate the forecasting and valuation equations after simultaneously imposing the constraint that the coefficient on growth in long-term net operating assets equals the coefficient on accruals in the forecasting equation ($\beta_1 = \beta_2$) and that the coefficients on growth in long-term net operating assets and accruals are also equal in the valuation equations ($\beta_1^* = \beta_2^*$). Panel B shows that, consistent with H2, the forecasting and valuation coefficients on GrLTNOA are equivalent to the forecasting and valuation coefficients on ACC ($p = 0.42$). We conclude that, after controlling for the mispricing of current ROA, the market's mispricing of growth in long-term net operating assets, relative to its implications for one-year-ahead ROA, does not differ significantly from the mispricing of accruals.

Supplemental Long-Run Abnormal Return Tests

We assess the robustness of our inference that the market overvalues accruals and growth in long-term net operating assets with a calendar-time portfolio regression described in Fama and French (1993) (hereafter Fama-French). Fama-French three-factor regressions on calendar-time portfolios are generally well specified for long-run abnormal return tests (Lyon et al. 1999).

To simulate a more realistic trading strategy, we restrict the analysis to calendar year-end firms. We form separate GrLTNOA and ACC portfolios of securities based on the highest and lowest deciles of either growth in long-term net operating assets (GrLTNOA) or accruals (ACC). We calculate returns for the 12 months beginning April following the fiscal year-end, when information to calculate growth in long-term net operating assets and accruals should be available from the firms' financial statements. We regress these adjusted returns for either GrLTNOA or ACC portfolios on market, firm size, and book-to-market factors as shown in Equation (4):

$$R_{pt} - R_{ft} = \alpha_p + b_p(R_{mt} - R_{ft}) + s_p(SMB_t) + h_p(HML_t) + \xi_{pt} \tag{4}$$

where R_{pt} is the monthly return on the GrLTNOA or ACC portfolio (value-weighted), R_{ft} is the one-month Treasury bill rate at the beginning of the month, R_{mt} is the value-weighted monthly return on all NYSE, AMEX, and NASDAQ stocks, SMB_t is the difference between returns of portfolios of small and large stocks (below or above the NYSE median, respectively), and HML_t is the difference between returns of high and low book-to-market stocks (above and below the 70 percent and 30 percent fractiles of book-to-market, respectively).[15] The factor loadings, b_p, s_p, and h_p, are the slopes of the three factors obtained by estimating the time-series regressions. The error term in the regression is denoted by ξ_{pt}. The estimate of the intercept, α_p, tests the null hypothesis that the mean monthly excess return on the calendar-time portfolio is zero, after controlling for the market, firm size, and book-to-market factors. Following Fama (1998), to control for variation in the number of securities in the portfolio over the sample period, we use a weighted-least-squares estimation, weighting portfolio returns by the number of securities included in the portfolio.

---

[14] This result also implies that the market underprices cash flows from operations relative to their implications for one-year-ahead ROA, after controlling for growth in net operating assets.

[15] We compute monthly measures of the three independent variables as in Fama and French (1996, see legend to their table 1).

Panel A of Table 5 reports the results for high GrLTNOA and ACC portfolios, and Panel B reports the results for low GrLTNOA and ACC portfolios. For both high-GrLTNOA and high-ACC portfolios, we find evidence of significant negative abnormal returns in the year after portfolio formation. Specifically, the estimated mean monthly abnormal returns (i.e., the intercepts) imply one-year-ahead annual abnormal returns of approximately −1.9

TABLE 5
Results of Alternative Tests of Long-Run Abnormal Stock Returns to Portfolios of High vs. Low Growth in Net Operating Assets

Panel A: Calendar-Time Portfolios Returns—High-Growth Portfolios

	HIGH GrLTNOA		*HIGH ACC*	
Fama-French Parameters	*Monthly Average*	*t-statistic*	*Monthly Average*	*t-statistic*
Intercept (α)	−0.0016	−1.70*	−0.0026	−2.67***
Market sensitivity (β)	1.1017	46.46***	1.1221	46.90***
Size sensitivity (s)	0.8933	25.03***	1.2683	35.21***
Book-to-market sensitivity (h)	0.0168	0.43	−0.0226	−0.57
Adjusted R^2	0.92		0.94	

Panel B: Calendar-Time Portfolios Returns—Low-Growth Portfolios

	LOW GrLTNOA		*LOW ACC*	
Fama-French Parameters	*Monthly Average*	*t-statistic*	*Monthly Average*	*t-statistic*
Intercept (α)	0.0045	3.22***	0.0053	4.62***
Market sensitivity (β)	1.0872	34.09***	1.0324	36.87***
Size sensitivity (s)	1.3766	28.64***	1.2519	29.76***
Book-to-market sensitivity (h)	0.5353	10.22***	0.3626	7.92***
Adjusted R^2	0.89		0.90	

***, **, and * Denote significance at the 1, 5, and 10 percent levels (two-tailed test), respectively.

We estimate abnormal returns using a cumulative procedure across the return interval, as shown in Fama and French (1993), for high- and low-growth portfolios, respectively. We form portfolios of securities based on the highest and lowest deciles of growth in long-term net operating assets (GrLTNOA) or operating accruals (ACC), and estimate the following Fama and French (1993) regression:

$$R_{pt} - R_{ft} = \alpha_p + b_p(R_{mt} - R_{ft}) + s_p(SMB_t) + h_p(HML_t) + \xi_{pt}, \quad (4)$$

where R_{pt} is the simple return on the calendar-time portfolio (value-weighted), R_{ft} is the monthly return on three-month Treasury bills, R_{mt} is the return on a value-weighted market index, SMB_t is the difference in the returns of portfolios of small and large firms, and HML_t is the difference in the returns of value-weighted portfolios of high book-to-market stocks and low book-to-market stocks.

percent and −3.1 percent to high-GrLTNOA and high-ACC portfolios, respectively. In contrast, for both low-GrLTNOA and low-ACC portfolios, we find evidence of significant positive abnormal returns in the year after portfolio formation: the monthly averages imply one-year-ahead annual abnormal returns of approximately 5.5 percent and 6.5 percent, respectively. Thus, our earlier findings that the market appears to overvalue both accruals and growth in long-term net operating assets are robust to controlling for the factors in the Fama-French three-factor asset-pricing model.

In addition, untabulated analyses reveal that the abnormal returns (i.e., the intercepts) do not differ significantly across the high-ACC vs. the high-GrLTNOA portfolios ($p > 0.80$), or across the low-ACC vs. the low-GrLTNOA portfolios ($p > 0.45$), using either nonparametric Wilcoxon rank sum tests or parametric tests of differences in portfolio means. Thus, for both the high-growth and low-growth portfolios, the mispricing of growth in long-term net operating assets does not significantly differ from the mispricing of accruals, after controlling for the factors in the Fama-French model.

V. CONCLUSIONS

Prior research finds differential persistence of the cash flows and accruals components of earnings for one-year-ahead ROA (e.g., Sloan 1996). However accruals is a component of both ROA and growth in net operating assets. We find that, after controlling for current ROA, both components of the growth in net operating assets—accruals and growth in long-term net operating assets—have equivalent negative associations with one-year-ahead ROA.

We interpret these results as suggesting that the lower persistence of accruals relative to cash flows documented in Sloan (1996) is a specific manifestation of a more general negative incremental relation between one-year-ahead ROA and growth in net operating assets. To the extent that the lower persistence of accruals arises because accruals are a component of growth in net operating assets, the lower persistence of accruals is more likely to result from the conservative bias in accounting or the lower rate of economic profits that result from diminishing marginal returns to new investment opportunities, or both, than from earnings management. These results are consistent with assertions that accruals models that do not control for growth may lead to inappropriate conclusions about earnings management (McNichols 2000).

Future research could investigate whether studies showing mispricing of abnormal accruals (Xie 2001; DeFond and Park 2001) have identified contexts in which earnings management is a central issue, or whether growth in net operating assets is a confounding factor that has been omitted from the research design. Research could also attempt to differentiate between balance sheet growth resulting from earnings management vs. balance sheet growth resulting from economic investment.

We also find that the well-documented anomaly wherein the market appears to overvalue accruals extends to the other component of growth in net operating assets—growth in long-term net operating assets. Thus, our results suggest that the accrual anomaly is likely attributable to the market mispricing of *growth in net operating assets*, regardless of whether the growth emerges in the form of changes in working capital or growth in long-term net operating assets. In sum, our evidence suggests that the accrual anomaly documented in Sloan (1996) may be one facet of a more general anomaly whereby the market misprices growth in net operating assets.

REFERENCES

Abarbanell, J., and B. Bushee. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research* 35 (Spring): 1–24.

Abel, A., and F. Mishkin. 1983. An integrated view of tests of rationality, market efficiency, and the short run neutrality of monetary policy. *Journal of Monetary Economics* 11 (January): 3–24.

Bahnson, P., P. Miller, and B. Budge. 1996. Nonarticulation in cash flow statements and implications for education, research, and practice. *Accounting Horizons* 10 (December): 1–15.

Barth, M., D. Cram, and K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (January): 27–58.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics*. New York, NY: Wiley.

Bernard, V., and J. Thomas. 1989. Post-earnings announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27 (Supplement): 1–36.

Collins, D., and P. Hribar. 2000. Earnings-based and accrual-based market anomalies: One effect or two? *Journal of Accounting and Economics* 29 (October): 101–123.

DeFond, M., and C. Park. 2001. The reversal of abnormal accruals and the market valuation of earnings surprises. *The Accounting Review* 76 (July): 375–404.

Drtina, R., and J. Largay. 1985. Pitfalls in calculating cash flows from operations. *The Accounting Review* 60 (April): 314–326.

Fairfield, P., R. Sweeney, and T. Yohn. 1996. Accounting classification and the predictive content of earnings. *The Accounting Review* 71 (July): 337–355.

———, and T. Yohn. 2001. Using asset turnover and profit margin to forecast changes in profitability. *Review of Accounting Studies* 6 (December): 371–385.

Fama, E., and J. MacBeth. 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 81 (May): 607–636.

———, and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (June): 427–465.

———, and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (February): 3–56.

———, and ———. 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51 (March): 3–56.

———. 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49 (September): 283–306.

Feltham, G., and J. Ohlson. 1995. Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research* 11 (Spring): 689–731.

Freeman, R., and S. Tse. 1989. The multiperiod information content of accounting earnings: Confirmations and contradictions of previous earnings reports. *Journal of Accounting Research* 27 (Supplement): 49–79.

Hribar, P., and D. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (March): 105–134.

Lakonishok, J., A. Shleifer, and R. Vishny. 1994. Contrarian investment, extrapolation, and risk. *Journal of Finance* 49 (December): 1541–1578.

Lyon, J., B. Barber, and C. Tsai. 1999. Improved methods for tests of long-run abnormal stock returns. *Journal of Finance* 54 (February): 165–201.

McNichols, M. 2000. Research design issue in earnings management studies. *Journal of Accounting and Public Policy* 19 (4–5): 313–345.

Mishkin, F. 1983. *A Rational Expectations Approach to Macroeconomics: Testing Policy Effectiveness and Efficient-Markets Models*. Chicago, IL: University of Chicago Press.

Ohlson, J. 1995. Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research* 11 (Spring): 661–87.

Ou, J., and S. Penman. 1989. Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics* 11 (November): 295–329.

———. 1990. The information content of nonearnings accounting numbers as earnings predictors. *Journal of Accounting Research* 28 (Spring): 144–163.

Penman. S. 2001. *Financial Statement Analysis and Equity Valuation*. Boston, MA: McGraw-Hill.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (July): 289–315.

Stigler, G. J. 1963. *Capital and Rates of Return in Manufacturing Industries*. Princeton, NJ: Princeton University Press.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76 (July): 357–373.

Editorial Data

The following table contains information about turnaround time for manuscripts (including revisions) on which editorial decisions were made in the 12-month period ended August 31, 2002. Turnaround time is the number of days between the date that the manuscript was received *and the date of the editor's letter to the author(s)*:

	Number of Manuscripts	Cumulative Number	Cumulative Percent
0 ≤ Days ≤ 30	81	81	22.50
31 ≤ Days ≤ 60	120	201	55.83
61 ≤ Days ≤ 90	115	316	87.78
91 ≤ Days ≤ 120	32	348	96.67
121 ≤ Days	12	360	100.00

The mean review time was 56 days; the median review time was 56 days.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal....No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of *The Chicago Manual of Style* (14th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be typed in 12-point font on one side of 8½ × 11" good quality paper and be double spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form*. See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings.

Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix a, b, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says...."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by *The Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Date of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2nd edition. Homewood, IL: Dow Jones-Irwin.
Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, U.K.: Cambridge University Press.
Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–636.
———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.
Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, The University of Texas at Austin.
Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan or the like) should be submitted. Information that might identify the author(s) must be deleted from the instrument.
3. The manuscript should be sent in an email as a Word or PDF attachment, with the title page in a separate file. ONE paper copy should be submitted with a check in U.S. funds for $75.00 for members or $100.00 for nonmembers of the AAA made payable to the American Accounting Association. The submission fee is nonrefundable. If preferable, the manuscript can be sent on a disk instead of as an email attachment, along with the one paper copy and the check.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting.

"An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

JANUARY 2003 PLACEMENT ADS

The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The **Review** *is published on the first day of January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding.*

101 UNIVERSITY OF NOTRE DAME, Department of Accountancy, invites applications for position openings at the Assistant and Associate Professor levels. All research areas will be considered. Applicants for Assistant Professor positions should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong commitment to scholarly research and excellence in teaching. Salary is very competitive and research support is excellent. Please send *curriculum vitae* and an example of scholarly work to Professor Tom Schaefer, Chair, Department of Accountancy, Room 102D, Mendoza College of Business Administration, University of Notre Dame, Notre Dame, IN 46556-5646. The University of Notre Dame is an Equal Opportunity, Affirmative Action Employer.

102 UNIVERSITY OF CALIFORNIA, IRVINE, Graduate School of Management, is seeking candidates for two tenure-track positions in Accounting to begin July 1, 2003. We anticipate filling one position at Associate or Full Professor rank and one at the Assistant, Associate, or Full Professor rank. Candidates from all areas of research and teaching will be considered. Candidates must have a Ph.D. or be in the final stages of dissertation completion, and have a strong commitment to both research and teaching, primarily at the graduate level. Candidates for senior positions should have a distinguished record of publications in top academic journals. Salaries are competitive. Applications received by January 31, 2003 will be considered. Applicants should submit their *curriculum vitae,* working papers, teaching evaluations, dissertation proposal, and letters of recommendation to Professor Barbara A. Lougee, Graduate School of Management, University of California, Irvine, CA 92697-3125. UCI is an Equal Opportunity Employer committed to excellence through diversity.

103 NICHOLLS STATE UNIVERSITY, Department of Accounting and Business Law, is seeking applicants for a tenure-track faculty position at the Assistant or Associate Professor level beginning Fall 2003. Primary duties include teaching undergraduate and graduate courses in Financial Accounting. Preferred candidates will have a Ph.D. in Accounting from an AACSB-accredited school at the time of appointment. Professional certification and experience are highly desirable. A strong record of scholarly publication and teaching experience is required for an appointment at the Associate Professor rank. Visit http://www.nicholls.edu/accounting for additional information. To apply, send application letter, current *curriculum vitae,* and three reference letters to Dr. Shawn Mauldin, Head, Department of Accounting and Business Law, PO Box 2015, Thibodaux, LA 70310. Review will begin immediately continue until the position is filled. NSU is an AA/EOE.

104 SIMON FRASER UNIVERSITY, the Faculty of Business Administration, invites applications for a tenure-track appointment in Financial Accounting or Managerial Accounting at the rank of Assistant/Associate Professor. Ph.D. in hand or near completion is required, and a professional Accounting designation preferred. Salary is competitive within Canada. Successful candidate will have strong research and teaching abilities with established research programs in one of the areas above and an interest in teaching at undergraduate/graduate/executive levels and in the Faculty's innovative online graduate programs. Start date is September 2003. Positions are subject to budgetary approval. Send *curriculum vitae* and the names of three referees to Appointments Committee Chair, Faculty of Business Administration, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6 CANADA. For more information, visit the following web sites: http://www.sfu.ca and http://www.bus.sfu.ca. All qualified candidates are encouraged to apply, however, Canadians and Permanent Residents will be given priority. SFU is committed to principle of equity in employment and offers equal employment opportunities to qualified applicants.

105 WICHITA STATE UNIVERSITY invites applications for the Omer Endowed Professorship in Accountancy and two tenure-eligible positions as Assistant/Associate Professor in Accountancy. To ensure full consideration applications must be received by November 1, 2002. The positions remain open until filled. Salary is competitive and the positions are contingent on funding. The position statements and application process are available at: http://webs.wichita.edu/accountancy/. An up-to-date *curriculum vitae* and names, addresses, and phone numbers of three professional references should be sent to Bill D. Jarnagin, Director, School of Accountancy, Wichita State University, 1845 Fairmount, Wichita, KS 67260-0087; Email: bill.jarnagin@wichita.edu; Phone: (316) 978-6262. Wichita State University is an Equal Opportunity/Affirmative Action Employer.

106 CLARKSON UNIVERSITY, School of Business, in Potsdam, New York, invites applications for a tenure-track position at the Assistant Professor level beginning in Fall 2003. The School of Business at Clarkson is AACSB-accredited, and offers Accounting courses at the undergraduate, M.S., and M.B.A. levels. Opportunities exist to teach courses in Financial and Managerial Accounting, Intermediate Accounting, Government and NFP Accounting, and Accounting Information Systems. Teaching load will be two sections per semester. Candidates must possess or be near completion of a Ph.D. in Accounting. Quality teaching and scholarship are necessary for tenure and promotion. Screening will begin immediately and continue until the position is filled. Submit a letter application, *curriculum vitae*, evidence of teaching effectiveness, a research sample, and three letters of recommendation to Chair, Accounting Search Committee, School of Business, Box 5765, Clarkson University, Potsdam, NY 13699; Email: suew@clarkson.edu. Clarkson University is an AAE/EOE. Pos #02-02.

107 POSITION DESIRED: Ph.D. is seeking tenure or tenure-track faculty position in the Accounting field. I have taught Accounting, Financial Accounting, Auditing, and AIS for more than 15 years at the University level. You may contact me by email at yugotit@empal.com.

108 FLORIDA INTERNATIONAL UNIVERSITY, one of the newest Florida university units, is seeking applicants for one Eminent Scholar research-oriented position starting Spring 2003, jointly in Accounting Information Systems and Management Information Systems. Closing date is 9/24/2002; Ph.D. or DBA is required. Salary is competitive and summer opportunities are available. FIU has a full-time faculty of about 20. Applicants should have extensive recent publications in leading refereed academic journals and proven ability to work with junior faculty. For consideration for an interview, please contact Dana A. Forgione, Chair, Search and Screen Committee, School of Accounting, College of Business Administration, Florida International University, Miami, FL 33199. Florida International University is an Equal Opportunity/Access Employer and Institution and member of the State University System of Florida.

109 UNIVERSITY OF CINCINNATI, Department of Accounting, invites applications for a tenure-track position at the Associate or Full Professor level beginning in Fall 2003. The qualifications include an established research record in leading academic journals and demonstrated success in teaching. Our primary research needs are in the empirical capital markets area, although other areas of specialization will be considered. The College of Business Administration provides access to most major financial databases through WRDS. The successful candidate will be required to teach in our Ph.D. program and take a leadership role in mentoring junior faculty research and Ph.D. students. The College of Business Administration is AACSB-accredited at the B.B.A. and M.B.A. levels and offers M.S. degrees in Accounting and Taxation. Salary and faculty support are competitive. Send a *curriculum vitae* and a working paper to Professor P. K. Sen, Department of Accounting, University of Cincinnati, Cincinnati, OH 45221-0211. The University of Cincinnati is an Affirmative Action/Equal Employment Opportunity Employer.

110 LONGWOOD UNIVERSITY, College of Business & Economics, an AACSB-accredited institution, seeks one individual at the Assistant Professor level for a tenure-track faculty position, beginning August 2003, teaching undergraduate Accounting. Candidates should possess an appropriate earned doctorate or be an ABD near completion. Professional certification is desirable. A demonstrated commitment to teaching excellence and an active research program is required. Salary is competitive and commensurate with qualifications and experience. Candidates must possess the willingness and ability to address core areas of the undergraduate Financial Accounting curriculum. Specializations and teaching interests include Auditing, Tax, or AIS. Candidates should submit a letter of interest; a current *curriculum vitae;* and the names, addresses, phone numbers, and email addresses of three references to Dr. Mary Flanigan, Chair, College of Business & Economics, Longwood University, 201 High Street, Farmville, Virginia 23909.

111 RICE UNIVERSITY, Jones Graduate School of Management, invites applications for tenure-track positions in Accounting at the Assistant and Associate Professor levels, beginning Fall 2003. Candidates with research interests in Empirical Financial Accounting or Empirical Managerial Accounting are preferred. Candidates for Assistant Professor should hold a Ph.D. or be nearing completion of their degree, and should demonstrate excellent research potential and evidence of teaching effectiveness. Candidates for Associate Professor should hold a Ph.D. and have a record of strong research productivity in leading journals, as well as evidence of high-quality teaching. Applicants should send *curriculum vitae,* three letters of recommendation, and recent working papers or dissertation proposal to Associate Dean Robert A. Westbrook, Jesse H. Jones Graduate School of Management—MS 531, Rice University, PO Box 1892, Houston, TX 77251-1892. Rice University is an Equal Opportunity/Affirmative Action Employer.

112 INDIANA UNIVERSITY, Department of Accounting and Information Systems, invites applications for a tenure-track faculty position effective Fall 2003. Preference will be given to Assistant Professors with teaching and research interests in Financial Accounting and Tax, but individuals at all ranks are welcome to apply. Candidates must possess a doctoral degree and have a strong commitment to excellence in scholarly research and teaching. Compensation levels are competitive and negotiable. Interested candidates should send a current *curriculum vitae,* research paper suitable for presentation at a workshop, and evidence of teaching ability. Three letters of reference will also be required before an offer of employment can be considered. Contact address is Jamie Pratt–Chair, Department of Accounting & Information Systems, KPMG Professor of Accounting, Kelley School of Business, 1309 E. 10th Street, Bloomington, IN 47405-1701; Email: jpratt@indiana.edu. Indiana University is an Equal Opportunity Affirmative Action Employer. Minorities, women, and persons with disabilities are encouraged to apply.

113 UNIVERSITY OF RICHMOND, Robins School of Business, invites applications for a tenure-track position, beginning August 2003. Applicants should possess a Ph.D./D.B.A. in Accounting or be in the final stages of completion. Candidates with teaching interests in Cost/Managerial Accounting and/or Accounting Information Systems are preferred. The University is a highly selective private university located in a residential area of Richmond, Virginia. The Robins School of Business and Accounting Department are AACSB-accredited. Please visit our web site at http://www.richmond.edu for more information. Applicants should send a resume to Professor Raymond Slaughter, Robins School of Business, University of Richmond, Richmond, VA 23173. The University of Richmond is an Equal Opportunity Employer.

115 MANHATTAN COLLEGE, School of Business, seeks applicants for a tenure-track position as an Assistant or Associate Professor of Accounting, beginning in August 2003. Candidates must have a Ph.D. or D.B.A. in Accounting and evidence of teaching effectiveness and research competence. Responsibilities include teaching undergraduate courses in Accounting, engaging in scholarly activities, and participating in campus activities. The Department seeks a candidate with a specialization in Taxation, but all qualified candidates are encouraged to apply. Recruiting begins October 1 and continues until the position is filled. Salary and rank will be based on qualifications. Please send a letter of application, *curriculum vitae,* and list of three references to Dr. James Suarez, Dean, School of Business, Manhattan College, Riverdale, NY 10471; Email: james.suarez@manhattan.edu. Manhattan College is an Affirmation Action/Equal Opportunity Employer.

116 UNIVERSITY OF SOUTH CAROLINA, School of Accounting, invites applications for a tenure-track position effective Fall 2003. We encourage applications from all ranks and areas of interest, but our preference is for an Assistant or Associate Professor. Applicants must possess a doctorate or expect to complete their doctorate by August 2003. In addition to being committed to excellent teaching, applicants must be ambitious and motivated researchers. The School of Accounting offers B.S., Master's of Accountancy, and Ph.D. degrees. We have a large number of faculty who publish in leading academic journals. We offer competitive salaries. We also look forward to hiring several new faculty members over the next few years. Please submit a resume and working papers or a defended dissertation proposal, if at the dissertation stage. Applications will be accepted until the position is filled. Submit applications to Professor Robert A. Leitch, The Moore School of Business, University of South Carolina, Columbia, SC 29208; Email: leitch@moore.sc.edu. The University of South Carolina is an Affirmative Action/Equal Opportunity Institution.

117 MIDDLE TENNESSEE STATE UNIVERSITY is accepting applications for four full-time, tenure-track positions (Assistant or Associate Professor) to begin August 1, 2003. Candidates should possess an earned doctorate (ABD will be considered if applicant is close to completion) in Accounting or a related area, though an L.L.M. or J.D. with relevant experience will be considered for applicants in appropriate areas. Preferred areas of teaching are Auditing, Tax, Systems, and Business Law. Preference will be given to candidates with the ability or willingness to incorporate innovative pedagogical methods. Phi Beta Kappa members are encouraged to apply. Review of applications will begin November 1, and will continue until positions are filled. Submit application letter, resume, transcripts, and three letters of reference to Dr. G. Robert Smith, Jr., Search Committee Chair, MTSU Box 50, Murfreesboro, TN 37132. Middle Tennessee State University is an EO/AA/AOA Employer. Applications from minorities and women are strongly encouraged.

118 SETON HALL UNIVERSITY, Department of Accounting, invites applications for two tenure-track positions beginning Fall 2003. All ranks will be considered. Applicants must have a primary teaching interest in either Financial Accounting or Auditing. Applicants must have a Ph.D. or equivalent. ABDs nearing completion will be considered. Applicants for the ranks of Associate and Full Professor must have an established and continuing record of accounting research. The Department offers a bachelor's degree in Accounting along with four different graduate degrees. The School of Business is AACSB-accredited. The Department has a strong emphasis on teaching excellence. Applicants should submit a *curriculum vitae,* copies of recent teaching evaluations, and three references. Applications should be directed to Dr. Karen Boroff, Dean, Stillman School of Business, Seton Hall University, South Orange, NJ 07079-2692; Email: Boroffka@shu.edu. Application screening will begin immediately and continue until positions are filled. Seton Hall University is an Affirmative Action/Equal Opportunity Employer.

119 THE AMERICAN UNIVERSITY IN CAIRO, Department of Management, invites applications for an Accounting teaching position. Rank is open; Ph.D. is required. Teaching/industry experience, professional certification, and scholarly research and publications are preferred. All areas of specialization will be considered. Responsibilities include graduate and undergraduate teaching, assistance with preparation for accreditation by AACSB, and scholarly research. One- to three-year appointments begin September 2003; possibility for renewal. Teaching load is three courses per semester. English is the language of instruction. Salary and rank are according to scale. For expatriates, housing and annual round-trip air travel are provided for appointee and accompanying family, as well as many other benefits; U.S. citizens preferred. Email application, specifying Position # MGMT-4 with *curriculum vitae* and names and addresses of three references, to facultyaffairs@aucnyo.edu or mail to Dr. Earl (Tim) Sullivan, Provost, The American University in Cairo, 420 Fifth Avenue, Fl. 3-AR, New York, NY 10018-2729, and complete the Personnel Information Form at http://forms.aucegypt.edu/provost/pif3.html. Applications will be accepted until position is filled. The American University in Cairo is an Equal Opportunity Employer.

120 UNIVERSITY OF PENNSYLVANIA invites applications from qualified persons for positions available July 2003 in Financial, Managerial, and Tax Accounting at all professional ranks. A Ph.D. is required. Demonstrated ability or potential for effective teaching and quality research are important considerations. Interested persons should send their resume including educational background, professional experience, publications, personal data, and, if applicable, a dissertation proposal and abstract to Professor Robert Holthausen, Accounting Department, University of Pennsylvania, Suite 2400, Steinberg-Dietrich Hall, Philadelphia, PA 19104-6365. The deadline for applications is March 2003. The University of Pennsylvania is an Equal Opportunity/Affirmative Action Employer and is strongly and actively committed to diversity within its community. Women and minority candidates are encouraged to apply.

121 THE UNIVERSITY OF TEXAS AT ARLINGTON, Department of Accounting, invites applications and nominations for the position of Department Chair at the level of Professor with tenure. Candidates should hold an earned doctorate, present evidence of successful administrative performance, and the ability to operate effectively in a collegial environment, demonstrate strong abilities in teaching and as a scholar, with well-documented research records and evidence of continuing research agendas. The University of Texas at Arlington, the second largest component in The University of Texas system, is a full-service research and teaching university with over 24,000 students. The College of Business Administration is fully AACSB-accredited and offers doctoral, master's, and undergraduate programs. The Department of Accounting offers a number of degrees including B.S., B.B.A., M.S.A., M.S.T., and M.P.A. Review of applications will begin December 13, 2002 and will continue until the position is filled. Please submit your material to Dr. Jeffrey E. McGee, Search Committee Chair, Department of Management, The University of Texas at Arlington, Box 19467, Arlington, Texas, 76019-0467. The University of Texas at Arlington is an Equal Opportunity/Affirmative Action Employer.

122 WHEATON COLLEGE seeks nominations for the William Volkman Chair of Business Law. Candidates for this tenure-track position should have experience in teaching Business and/or Economics and Law. The holder of the Volkman Chair will assume the leadership of the College's Prelaw Studies Program where teaching and advising students in preparation for admission to law school is expected. The candidate should hold a J.D. and preferably another academic or professional credential (e.g., Ph.D., D.B.A., M.B.A., CFA, or CPA). Nominations and inquiries should be directed to Dr. Dorothy F. Chappell, Dean of Natural and Social Sciences, Wheaton College, 501 College Avenue, Wheaton, IL 60187, no later than November 1, 2002. Wheaton College is an evangelical Christian liberal arts college whose faculty and staff affirm a Statement of Faith and adhere to lifestyle expectations. The College complies with federal and state guidelines for nondiscrimination in employment. Women and minorities are encouraged to apply.

123 PORTLAND STATE UNIVERSITY, School of Business Administration, invites nominations and applications for the Retzlaff Chair in Accounting. This tenure-track position will be a thought leader, pursuing research that adds value to organizations by addressing issues associated with the internal use of information. The successful candidate will address and promote professionalism in Accounting and the role of Accounting in managing and sustaining successful organizations. The candidate should have an established national reputation in Accounting research and the capability to provide school-wide leadership. Rank is subject to applicant's qualifications. Preference will be given to applicants whose research is compatible with that of existing faculty. The successful applicant will be expected to provide graduate and undergraduate teaching support within the current curriculum. A commitment to community service appropriate for an urban university is expected. To apply, contact Darrell Brown at (503) 725-5099 or by email at darrellb@sba.pdx.edu. Portland State University is an Affirmative Action, Equal Opportunity Institution.

124 NATIONAL UNIVERSITY OF SINGAPORE (NUS), Department of Finance and Accounting, invites applications for accounting tenure-track positions at all ranks. All accounting-related areas will be considered. Candidates should possess a Ph.D. degree or be expected to obtain one soon. Selection for Associate or Full Professor ranks will be based on an established publication record in reputable journals and proven teaching excellence. Selection for the Assistant Professor rank will be based on potential for publications in reputable journals and for effective teaching. Applications with full resume and three letters of recommendation should be sent to Associate Professor Trevor Wilkins, Head of Department, Department of Finance and Accounting, National University of Singapore, 1 Business Link, Singapore 117592; Email: fnbhead@nus.edu.sg.

125 RUTGERS UNIVERSITY, Camden, School of Business, invites applications for a tenure-track faculty position. The rank is open. The candidate should be capable of high-quality research and teaching at both the undergraduate and graduate levels. The school is AACSB-accredited. The normal teaching load is two courses (sections) per semester. Salary is competitive and commensurate with qualifications. A doctorate in Accounting is required, but ABD will be considered. Additional CPA certification is preferred. The University is located in southern New Jersey and is less than two miles from downtown Philadelphia, with easy access to New York and Washington, D.C. The position will remain open until filled. Interested candidates should send resumes to Emel Kahya, The Accounting Area Coordinator, Rutgers University, School of Business, 227 Penn Street, Camden, NJ 08102; Phone: (856) 225-6724. Rutgers University values diversity and encourages women and ethnic minorities to apply. Rutgers University is an Affirmative Action/Equal Opportunity Employer.

126 MIDWESTERN STATE UNIVERSITY, Department of Accounting, invites applications for a tenure-track faculty position effective Fall 2003. A doctorate in Accounting and a record of teaching excellence, scholarly research, and publication are required. Appointment will be as an Assistant or Associate Professor, depending on qualification. Compensation is competitive and negotiable. Interested candidates should send a letter of application, current *curriculum vitae,* unofficial transcripts, and a list of three references to Dr. Ralph Fritzsch, Chair, Department of Accounting, College of Business Administration, Midwestern State University, 3410 Taft, Wichita Falls, TX 76308; Email: ralph.fritzsch@mwsu.edu. For additional information, please see: http://business.mwsu.edu/accounting/index.asp.

127 CALIFORNIA STATE UNIVERSITY, Stanislaus has a tenure-track Accounting position at the Assistant level beginning Fall 2003. Desired area of specialization is Information Systems. Individuals with a primary interest in Information Systems, recent professional experience, and professional certification are especially encouraged to apply. Excellent teaching, research productivity, and service to the institution and profession are expected for promotion and tenure. A Ph.D. in Accounting from an AACSB-accredited program and graduate coursework in Information Systems are required. Salary is competitive, commensurate with qualifications and experience. Interested persons should send a letter of application and *curriculum vitae* to the Search Committee, Department of Accounting and Finance, College of Business Administration, California State University, Stanislaus, 801 W. Monte Vista Avenue, Turlock, CA 95382. Screening will begin immediately, and will be open until filled. California State University, Stanislaus is an Equal Opportunity/Affirmative Action Employer.

128 PURDUE UNIVERSITY, Krannert Graduate School of Management, invites applicants for a tenure-track appointment in Accounting beginning Fall 2003. Rank is at either the Assistant or Associate Professor level. Candidates for an Assistant Professor position should have (or be near completion of) a doctorate and exhibit the capacity for research publishable in top-tier accounting journals and excellence in teaching. Candidates for an Associate Professor position should have an earned doctorate and an established record of excellence in research, teaching, and service on doctoral dissertation committees. We offer a competitive compensation package. Applications will be reviewed beginning September 15, 2002. Applications should include a resume, representative research output, evidence of teaching ability, and at least three letters of reference, and should be sent to Professor Mark Bagnoli, Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907. Purdue is an Affirmative Action, Equal Opportunity Employer. Women, minorities and individuals with disabilities are encouraged to apply.

129 UNIVERSITY OF MISSOURI–COLUMBIA (MU), School of Accountancy, seeks to fill the position of Director of the School beginning Fall 2003. Screening begins December 1, 2002, and will continue until the position is filled. We seek a proven leader to sustain and enhance our strong national prominence. Responsibilities include administration of the School's activities, fund-raising, and relations with stakeholders. Applicants must have a doctorate in Accounting and fulfill the qualifications for initial appointment as Full Professor, or have a demonstrated record of research and teaching that assures early promotion to Full Professor. Previous administrative experience is preferred. Interested candidates should visit our web site, http://business.missouri.edu/About/Academic/Acct/ and click on Director Position Opening for more information about the application portfolio. Application portfolios or inquiries should be sent to Dr. Loren Nikolai, School of Accountancy, 303 Cornell Hall, University of Missouri, Columbia, MO 65211; Email: nikolai@missouri.edu. MU is an Equal Opportunity/Affirmative Action Employer.

130 UTAH STATE UNIVERSITY, School of Accountancy, has a tenure-track faculty position available. Responsibilities include teaching at both undergraduate and master's levels in one or more of the following areas: Financial Accounting, Auditing, and Information Systems, publishing in nationally recognized journals, participation in academic and professional organizations, and involvement in university service. Requires Doctorate in Accounting. ABD candidates considered, but dissertation must be completed prior to appointment. See http://personnel.usu.edu (2-129) College of Business for full description. Send letter of interest and detailed *curriculum vitae* to Dr. I. Richard Johnson, Chair, Faculty Search Committee, 3540 Old Main Hill, Logan, UT 84322-3540. Review begins January 15, 2003 and will remain open until filled. AA/EOE.

131 GEORGIA COLLEGE & STATE UNIVERSITY seeks applicants for Dean of School of Business beginning August 1, 2003. The Dean reports directly to the Vice President and Dean of Faculties and is responsible for strategic planning, management, and development of the School of Business. Qualifications include an earned doctorate; teaching, research, and service profile commensurate with appointment to the rank of Professor; excellent communication and interpersonal skills; academic leadership/administrative experience; experience with AACSB accreditation; understanding of GC&SU's public liberal arts mission; willingness to seek external funding; ability to cultivate ties to alumni and area businesses; and a commitment to student achievement and faculty/staff development. Screening of applications and nominations will begin on January 2, 2003 and will continue until a suitable candidate is found. Submit letter of application, *curriculum vitae,* and five references to Dr. Mike Whitfield, Chair, Dean Search Committee, CBX 011, GC&SU, Milledgeville, GA 31061. Electronic applications and nominations may be submitted by email to mwhitfie@gcsu.edu. Georgia College & State University is an Equal Opportunity/Affirmative Action Employer.

132 KENT STATE UNIVERSITY, Department of Accounting, invites applications for an Assistant Professor with teaching and research emphasis in Managerial Accounting beginning in Fall 2003. A Ph.D. with a major in Accounting or an appropriate related field is required; candidates who are ABD will be considered. Professional certification such as CPA or CMA is highly desirable. Evidence of the potential for effective undergraduate and graduate teaching, and discovery/integration research is required. Research will be expected in high-quality, academic refereed journals for reappointment, tenure, and promotion; some publication in other journals may be included in the faculty member's portfolio. Participation in teaching doctoral students and serving on and/or supervising their dissertations likely will be part of this faculty member's responsibilities. Salary range is to $90,000; two summers of research support possible; summer teaching available. Send your application documents to Dr. Pervaiz Alam, Chair, Search Committee, Department of Accounting, 576 BSA, Kent State University, Kent, OH 44242; Phone: (330) 672-1121; Fax: (330) 672-2548; Email: palam@bsa3.kent.edu.

133 LOUISIANA STATE UNIVERSITY, Department of Accounting in the E. J. Ourso College of Business Administration, invites applications for one or more tenure-track openings to begin August 2003. Required Qualifications for Assistant/Associate Professor: will be considered in all disciplines, with priority given to Cost/Managerial or Tax specialties. Required Qualifications for Assistant Professor: Doctoral degree in Accounting be completed, or near completion, by the date of appointment (if degree is not completed by date of appointment, rank will initially be at Instructor level). Additional Qualification Desired: CPA. Responsibilities: teach undergraduate and graduate courses; be actively involved in academic research; provide service to the profession and the university. Applications will be accepted until March 1, 2003, or until filled. Please send a letter of application, *curriculum vitae* (including email address), and references to Don W. Finn, Chair, Department of Accounting, 3101 CEBA Building, LSU, Ref: Log #0408, Baton Rouge, LA 70803. LSU is an EO/AA Employer.

134 UNIVERSITY OF NORTH CAROLINA AT GREENSBORO, Bryan School of Business and Economics, Department of Accounting, invites applications for anticipated positions at the Assistant Professor and/or Associate Professor levels. Teaching and research interests in Auditing, Financial Accounting, and/or Tax Accounting with a Ph.D. or ABD status are required. Interested candidates should submit a letter of application, *curriculum vitae,* and the names of three references to Professor William A. Collins, Chair, Search Committee, Bryan School of Business and Economics, UNCG, PO Box 26165, Greensboro, NC 27402-6165. The University of North Carolina at Greensboro is committed to equality of employment opportunity and does not discriminate against applicants or employees based on race, color, national origin, religion, gender, age, disability, or veteran status. Moreover, the University of North Carolina at Greensboro is committed to recruiting and advancing women and minorities at all faculty/staff levels.

135 UNIVERSITY OF NORTH CAROLINA AT GREENSBORO, Bryan School of Business and Economics, Department of Accounting, invites applications for an anticipated opening for an Assistant Professor. Applicants should hold a Ph.D. or DBA or be in the final dissertation stage. Applicants must have a commitment to teaching excellence and the demonstrated interest and ability to do high-quality publishable research. The Department offers a B.S. with an Accounting major, a Master's of Science degree program in Accounting, and service courses for an M.B.A. program. Salary is competitive. Interested candidates should submit a letter of application, *curriculum vitae,* and the names of three references to Professor William A. Collins, Chair, Search Committee, Bryan School of Business and Economics, UNCG, PO Box 26165, Greensboro, NC 27402-6165. EEO/AA: W/M/V/D

136 MORGAN STATE UNIVERSITY, the Graves School of Business and Management, invites applications for a tenure-track position in Accounting at the Associate Professor or Professor rank. Candidates should have a Ph.D. degree in Accounting from an AACSB-accredited institution and a significant record of publications (Financial, Managerial, Auditing) in recognized Accounting journals. Salaries are AACSB-competitive. The Graves School offers a B.S. degree in Accounting, an M.B.A. degree, and a Ph.D. degree in Business Administration with a concentration in Accounting. The school has both Business and Accounting accreditation by the AACSB International. Send letter of interest and academic *vitae* to Dr. Sharon Gary Finney, Chairperson, Department of Accounting & Finance, Morgan State University, Baltimore, MD 21251, or email same to sfinney@moac.morgan.edu. Morgan State University is an Affirmative Action/Equal Opportunity Employer.

137 SOUTHERN CONNECTICUT STATE UNIVERSITY, Accounting Department, invites applications for a tenure-track faculty position. Rank will depend upon qualifications. Anticipated start date is Fall 2003. Responsibilities: graduate and undergraduate teaching, participating in preparation for AACSB accreditation and scholarly research. Qualifications: Ph.D. in Accounting or ABD nearing completion. Teaching/industry experience, professional certification, recent scholarly publication, and holding master's and doctoral degrees from AACSB-accredited institutions, are preferred. CPA/Master's qualification will be considered at the Assistant level, but earning a Ph.D. will be required for tenure. Applicants with the ability to teach Taxation and Principles of Financial or Managerial Accounting are preferred. Search process begins December 1, 2002 and will continue until position is filled. Candidates should submit application letter, *curriculum vitae,* and three references to Dr. Janet F. Phillips, Chair, Search Committee, Department of Accounting, Southern Connecticut State University, 501 Crescent Street, New Haven, CT 06515. SCSU is an Affirmative Action/Equal Employment Opportunity Employer.

138 NORTHEASTERN ILLINOIS UNIVERSITY, College of Business and Management, invites applications for Ph.D./D.B.A. with specialization in Accounting from an AACSB-accredited College or Business program in all or some of the following areas: Auditing, Management Consulting, International Accounting, and Multinational Business Operations. One tenure-track position at the Assistant/Associate Professor rank, beginning August 2003, is anticipated for teaching undergraduate and graduate courses. Potential for excellent teaching and research. Selection process will continue until position is filled. Send application letter, resume with email address, transcripts, and names of three references to Dr. C. T. Chen, Chair, Department of Accounting, Business Law and Finance, Northeastern Illinois University, 5500 N. St. Louis Avenue, Chicago, IL 60625; Fax: (773) 442-6110; Email: C-Chen3@neiu.edu. AA/EOE.

139 KEENE STATE COLLEGE, Management Department, invites applicants for an Assistant Professor tenure-track position in Accounting and MIS beginning Fall 2003. Earned doctorate by August 2003 from an accredited college or university, interest in teaching Quantitative Decision Making desirable; teaching experience is required. The Management Department emphasizes teaching excellence. The Department's mission focuses on preparing its graduates with a strong liberal arts background along with the management skills required to solve complex problems in an ever increasingly competitive global environment. Submit letter of interest, *curriculum vitae,* official college transcripts, three reference letters, recent teaching evaluations, evidence of scholarly work to Management Search, Keene State College, 229 Main Street, Keene, NH, 03435-1604. Review of application begins immediately and continues until position is filled. To learn more about Keene State College and Keene, New Hampshire, visit these web sites: http://www.keene.edu/ or http://www.KeeneNH.com/. Keene State College is an AA/EOE engaged in an effort to build a community that reflects the diversity of society.

140 UNIVERSITY OF WISCONSIN–PARKSIDE, School of Business and Technology, has an Assistant Professor position available. Doctorate in Accounting or J.D. with L.L.M. is required; ABD candidates considered. Teaching interests should be in the following areas: Auditing, Systems, Government, Tax, Not-for-Profit, and Legal Environment. Applicants must provide evidence of a strong commitment to quality teaching and scholarly research. Applicants, also, must have an interest in developing and maintaining successful interactions with the local community. Send application letter, current *curriculum vitae,* supporting evidence, and the names, addresses, and phone numbers of three persons who are willing to serve as references to Professor Jamie Wang, School of Business and Technology, University of Wisconsin–Parkside, 900 Wood Rd., Box 2000, Kenosha, WI 53141-2000. Review of applications begins immediately and will continue until position is filled. The School's undergraduate and M.B.A. programs are accredited by the AACSB. The University of Wisconsin–Parkside is an AA/EEO Employer D/M/V/W. Visit http://www.uwp.edu/academic/business/faculty/acctPosition.htm for more information.

141 TRUMAN STATE UNIVERSITY invites applicants for a tenure-track faculty position in Tax Accounting, rank commensurate with experience. Requirements: Candidates must meet one of the following criteria: (1) Ph.D. or D.B.A. in Accounting—ABDs nearing completion will be considered; (2) master's degree in Accounting or Taxation, a CPA license, and significant professional experience in Taxation; (3) Juris Doctorate with a master's degree in Taxation or Accounting. In addition, all candidates must have superior undergraduate credentials and must be supportive of a liberal arts and sciences curriculum. Send a letter of application, resume, three recent letters of recommendation, graduate and undergraduate transcripts to Dr. James Bailey, Division of Business and Accountancy, 100 East Normal, Kirksville, MO 63501. Position announcement located at http://business.truman.edu. EEO/AA/ADA.

142 MISSISSIPPI STATE UNIVERSITY, College of Business and Industry, School of Accountancy, is accepting applications for a tenure-track position of Assistant Professor of Accounting beginning Fall 2003. Responsibilities will include teaching undergraduate and graduate courses in either Accounting Information Systems, Auditing, and/or Managerial Accounting. Doctorate degree in Accounting is required. ABDs with dissertation near completion may be considered at instructor level. Requires excellence in teaching, based on prior experience as a graduate teaching assistant or as an instructor or Assistant Professor, ability to do research leading to publication in nationally refereed journals, and an interest and willingness to be involved in professional Accounting activities. An interest in advising and mentoring students is expected. Professional certification and professional Accounting experience is highly desirable. Salary is competitive. Send letter of application, current *curriculum vitae,* and three letters of reference to Dr. Dan Hollingsworth, Director, School of Accountancy, Drawer EF, Mississippi State, MS 39762. Application will remain open until a suitable candidate is found. MSU is an AAEOE.

143 UNIVERSITY OF MICHIGAN, Business School, has tenure-track positions available in Accounting starting in the 2003–2004 academic year. The application should include a research paper based on the candidate's dissertation, any other published or unpublished research, evidence of teaching experience (if any), and a *curriculum vitae* that includes three references. Applicants should have primary research and teaching interests in Financial or Managerial Accounting. Applications should be sent to Professor Russell J Lundholm, University of Michigan Business School, 701 Tappan Street, Ann Arbor, MI 48109-1234. The University of Michigan offers competitive salary and fringe benefits and is an Affirmative Action/Nondiscriminatory Employer.

144 CORNERSTONE RESEARCH, an economic consulting firm specializing in analyzing economic, financial, accounting, and marketing issues, is interested in hiring full-time associates for its offices in Los Angeles, Menlo Park, CA; Boston, MA; New York, NY; and Washington, D.C. Associates work directly with esteemed faculty experts in a distinctive "partnership" that combines the strengths of the academic and business worlds. We seek Ph.D. candidates with the ability to apply academic research to real-world issues, present concise explanations of complex analyses, and independently manage projects and junior staff. The ideal candidate will possess a strong empirical background in Financial, Cost, Managerial Accounting, or Auditing, as well as demonstrate excellent interpersonal skills. We are also seeking experienced candidates for more senior levels. Compensation commensurate with experience. To apply please send a cover letter that includes a statement describing your interest in economic consulting and your location preferences, resume, and two letters of reference. Contact Associate Recruiting Coordinator—Accounting, Cornerstone Research, 1000 El Camino Real, Menlo Park, CA 94025; web site: http://www.cornerstone.com. Cornerstone Research is an Equal Opportunity Employer.

145 LONG ISLAND UNIVERSITY, BROOKLYN CAMPUS, School of Business, Public Administration and Information Sciences, is growing rapidly. A tenure-track position in Accounting is open for Fall 2003 at the Assistant/Associate Professor level. Expertise in AIS and Managerial Accounting or Financial Accounting is required. Candidates must have an earned doctorate (or be ABD nearing completion); CPA or CMA preferred. Professional experience, certification, and teaching experience are required for senior placement. The Brooklyn Campus, occupying a ten-acre site in downtown Brooklyn, is one of the country's most dynamic and diverse urban campuses, serving students in day, evening, and weekend programs. It enrolls over 10,000 students in over 52 graduate and 62 undergraduate degree programs. Long Island University is the nation's seventh largest private university with six campuses in metropolitan New York and Long Island. Send letter of interest, *curriculum vitae,* and three reference contacts to Professor Frederic M. Stiner, Chair, Department of Accounting, Taxation and Law, Long Island University, 1 University Plaza, Brooklyn, New York 11201-8423; Fax: 718-488-1125; or Email: Frederic.Stiner@liu.edu. Search will remain open until position filled. AA/EOE.

香港浸會大學
HONG KONG BAPTIST UNIVERSITY

Department of Accountancy and Law
Professor/Associate Professor/Assistant Professor in Accounting
(PR006/2003)

The appointee is expected to teach Accounting subjects and supervise student projects at both the undergraduate and postgraduate levels. The candidate is also expected to conduct research in his/her own field and to perform administrative duties in addition to teaching/research. Applicants should possess a relevant PhD degree or equivalent professional standing with experience in teaching and research. A record of excellence in teaching at the tertiary level, and strong research background and leadership ability are expected for an appointment to senior rank. Knowledge of local accounting practice, and professional experience in industry and consultancy are desirable.

Salary Scales: (approximately US$1=HK$7.8)
Professor : HK$78,230 - HK$99,815 p.m.
Associate Professor : HK$72,020 - HK$87,505 p.m. (Scale A)
HK$61,940 - HK$77,435 p.m. (Scale B)
Assistant Professor : HK$46,355 - HK$68,145 p.m.

Terms of Appointment:
The appointee will initially be offered a fixed-term contract of three years commencing September 2003. Benefits include contribution by the University to a retirement benefits scheme and/or a gratuity payable upon satisfactory completion of contact (totalling up to 15% of basic salary), annual leave, medical & dental benefits for appointee & family, housing and relocation allowance where applicable. Re-appointment thereafter is subject to mutual agreement.

Application Procedure:
Application, together with updated curriculum vitae, copies of transcripts/testimonials should be sent to the Personnel Office, Hong Kong Baptist University, Kowloon Tong, Hong Kong [Fax: (852) 3411-5001; E-mail: recruit@hkbu.edu.hk]. Application forms can be obtained by downloading from: [http://www.hkbu.edu.hk/~pers]. Applicants are requested to send in samples of publications, preferably three best ones out of their most recent publications. Applicants should also request two referees to send in confidential reference to the Personnel Office direct. *Please quote PR number on the application, confidential reference letters and any subsequent correspondence.* General information about the University is available on the Web at [http://www.hkbu.edu.hk]. Details of the University's Personal Information Collection Statement can be found at [http://www.hkbu.edu.hk/~pers/job].

Closing date: 31 January 2003
The University reserves the right not to make an appointment for the posts advertised, and the appointment will be made according to the terms & conditions then applicable at the time of offer.

香港城市大學
City University
of Hong Kong

The City University of Hong Kong (CityU) is a leading university in the Asia Pacific region. CityU is located in close proximity to the hub of all business activities in Hong Kong with convenient access through the mass transit railway. The University has a student population of about 18,500 and offers excellent facilities with generous teaching and research support. CityU is committed to excellence in teaching and research as well as to maintaining close ties with commerce and industry. The medium of instruction is English given that Hong Kong has a significant role to play as an international city. The Faculty of Business is proud to be one of the largest business schools in Hong Kong with a total of about 180 faculty members within its six academic departments. All six departments are well known in the business community for their professional standing and academic reputation.

The Department of Accountancy is one of the six departments in the Faculty of Business. It has 29 full time faculty members and 6 administrative staff. One honorary research professor, namely **Professor Ross Watts**, William H Meckling Professor of Business Administration, William E Simon Graduate School of Business Administration, University of Rochester has been appointed for a period of three years to augment our research profile. Our department provides state-of-the-art accountancy education with a professional emphasis. We offer three undergraduate programmes namely the Bachelor of Business Administration (Honours) in Accountancy, Bachelor of Business Administration (Honours) in **Accountancy and Management Information Systems**, Bachelor of Business Administration (Honours) in **Accountancy and Law** and three taught postgraduate programmes including Master of Arts in International Accounting, Master of Arts in Professional Accounting and Information Systems and Postgraduate Certificate in Professional Accounting (PGCPA). Our undergraduate degrees and the PGCPA have been fully accredited by the Hong Kong Society of Accountants for their new Qualification Programme.

Research is strongly and actively encouraged and we have been successful in obtaining numerous competitive research grants both within and outside the University. The Department is further strengthening its PhD programme with a focus on economic aspects of accounting and auditing and corporate governance issues, with special interest in the Chinese Mainland and the Asia-Pacific region. Research seminars with eminent professors from the top US universities have become part of our regular activities.

Our Department has successfully established the interdisciplinary ***Accounting & Corporate Governance Centre***, headed by **Professor Ferdinand A Gul, Head & Professor (Chair) of the Department of Accountancy.** The mission of this Centre is to identify key issues and problems and undertake collaborative research in the areas of Accountancy and Corporate Governance to provide solutions and prescriptions for policy makers in Hong Kong, the Chinese Mainland and the region. It has established links with academic institutions in the United States, United Kingdom, Canada, Australia and other countries in the region with a view to conducting comparative studies in accounting and corporate governance. The Centre publishes "The Asia Pacific Journal of Accounting & Economics (APJAE)" with Professor Ferdinand A Gul and Professor Dan Simunic as Co-editors and Professor Katherine Schipper and Professor Ross Watts as consulting editors. Our excellent track record in internationally refereed publications such as ***The Accounting Review, Journal of Accounting and Economics, Contemporary Accounting Research, Accounting, Organizations and Society, Auditing: A Journal of Practice and Theory, Journal of Corporate Finance*** are evidence of our research strength.

Posts and Subject Areas

The Department is now inviting applications for the positions of **Professor/Associate Professor/Assistant Professor tenable from September 2003/January 2004.** The appointees will teach in one or more of the following areas: Financial Accounting / Accounting Information Systems / Auditing / International Accounting. Preference will be given to those with research and teaching interests in international and economic dimensions of financial accounting, auditing and the application of information technology to all areas of accounting. Our teaching load is normally two courses per semester. All undergraduate classes are conducted in the day time, with evening teaching for our taught postgraduate programmes. There will also be some additional undergraduate or postgraduate supervision of projects and thesis respectively.

The University also accepts nominations for the post of Professor and reserves the right not to fill the position.

General Requirements for Appointment

Applicants for Professor/Associate Professor should possess a **PhD**, a solid publication record in good international journals and a dedication to quality teaching. They are expected to provide contribution/leadership in research, plan curriculum of studies and supervise postgraduate students. They are required to have relevant teaching, and preference will be given to those with professional qualifications and research experience with a record of publications. Applicants for Assistant Professor should possess a PhD.

Salary # and Conditions of Service

Professor	approx. US$115,500 - $153,500 per annum
Associate Professor	approx. US$110,800 - $148,800 per annum
Assistant Professor*	approx. US$71,300 - $119,100 per annum

*Appointees at the level of Assistant Professor with salary at US$100,000 per annum or above may be considered for the title of Associate Professor.
#Standard income tax rate is 15%.

Appointment will be either on a three-year contract with contract-end gratuity or on superannuable terms with an initial evaluation period of 3 years. Fringe benefits include annual leave, medical and dental schemes, and housing assistance where applicable.

Information and Application

Information concerning the posts and the University is available on the University's World Wide Web home page http://www.cityu.edu.hk or the University's listserver accessed by E-mail at "hrmail@ctylnk.cityu.edu.hk" or from the Human Resources Office, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong (Fax : (852) 2788 1154 or 2788 9334 / E-mail : hrrecrut@cityu.edu.hk). Further information can be obtained from Professor Ferdinand A Gul, Head of Department of Accountancy at Personal Fax No. (852) 2788 7002 or E-mail: achead@cityu.edu.hk. Please send your nomination or application in the form of an application letter enclosing a current curriculum vita to the Human Resources Office.

澳門大學
Universidade de Macau
University of Macau

Academic Positions at the Faculty of Business Administration

The University of Macau, a government-funded institution located in the Macao Special Administrative Region of the People's Republic of China, invites applications for the academic positions at the assistant/associate/full professor levels in the following areas:

Accounting ***(ref.FBA/PROF/AC/2003),*** Financial Accounting, Managerial Accounting,
Accounting Information Systems, China Accounting and Taxation,
International Accounting, Auditing

Applicants should possess a relevant Ph.D. degree from an accredited institution. For the posts of assistant professor, Ph.D. candidates who expect to complete their studies prior to the starting date of appointment will also be considered. Applicants for posts at the associate professor level and above should have at least five years' full-time teaching record or other relevant professional experience, and a demonstrated ability to perform scholarly research.

Appointees are required to teach courses at the undergraduate and postgraduate levels and to provide MBA dissertation supervision. Appointments are expected to begin in September 2003.

Annual Salary
The salary level, shall be determined according to the appointee's academic qualifications, current position, and professional experience.

Assistant Professor:	MOP455,000 - MOP525,000
Associate Professor:	MOP560,000 - MOP630,000
Full Professor:	MOP644,000 - MOP700,000

(USD1 approx. = MOP8)

Additional information on terms of appointment for academic staff can be found on http://www.umac.mo/vacancies/terms-gnl.html

Application Procedure
Candidates should send a letter of application with C.V., copies of certificates, copies of IDs or passports, two letters of reference, quoting respective reference number to: The Personnel Office, The University of Macau, Av. Padre Tomas Pereira S.J.,Taipa, Macau

Website: http://www.umac.mo/vacancies/index.html
E-mail: PO.vacancy@umac.mo
Telephone: +853 3974 844
Fax no.: +853 3974 845 / +853 831694

NYU STERN

NEW YORK UNIVERSITY
LEONARD N. STERN SCHOOL OF BUSINESS
DEPARTMENT OF ACCOUNTING

NYU Stern School of Business invites highly qualified applications for tenure-track faculty positions at the Assistant Professor level commencing Fall 2003. The qualifications required include a demonstrated interest in and capacity to do publishable research, and a high level of teaching competence in Undergraduate, M.B.A. and Ph.D. classes. Candidates are expected to hold a doctorate degree or be assured of its completion within one year. All accounting-related discipline areas will be considered.

Send a resume and a list of references by January 31, 2003 to:

Professor James Ohlson, Acting Chair
New York University
Leonard N. Stern School of Business
40 West 4th Street, Suite 433
New York, NY 10012

New York University is an affirmative action / equal opportunity employer.

THE ACCOUNTING REVIEW
ISSN 0001-4826

PERIODICALS POSTAGE PAID
SARASOTA FLORIDA
AND ADDITIONAL MAILING OFFICES

POSTMASTER
Send Address Changes To:

AMERICAN ACCOUNTING ASSOCIATION
5717 BESSIE DRIVE
SARASOTA, FL 34233

THE ACCOUNTING REVIEW

18.07.03

VOLUME 78, No. 2 APRIL 2003

ASHIQ ALI, LEE-SEOK HWANG, and MARK A. TROMBLEY
Residual-Income-Based Valuation Predicts Future Stock Returns: Evidence on Mispricing vs. Risk Explanations

PETER F. CHEN and GUOCHANG ZHANG
Heterogeneous Investment Opportunities in Multiple-Segment Firms and the Incremental Value Relevance of Segment Accounting Data

ANDREW FERGUSON, JERE R. FRANCIS, and DONALD J. STOKES
The Effects of Firm-Wide and Office-Level Industry Expertise on Audit Pricing

BJORN N. JORGENSEN and MICHAEL T. KIRSCHENHEITER
Discretionary Risk Disclosures

KRISHNA R. KUMAR and GNANAKUMAR VISVANATHAN
The Information Content of the Deferred Tax Valuation Allowance

JOHN PHILLIPS, MORTON PINCUS, and SONJA OLHOFT REGO
Earnings Management: New Evidence Based on Deferred Tax Expense

STEPHEN G. RYAN and PAUL A. ZAROWIN
Why Has the Contemporaneous Linear Returns-Earnings Relation Declined?

KAREN L. SEDATOLE
The Effect of Measurement Alternatives on a Nonfinancial Quality Measure's Forward-Looking Properties

HUN-TONG TAN and KEN T. TROTMAN
Reviewers' Responses to Anticipated Stylization Attempts by Preparers of Audit Workpapers

---

A QUARTERLY JOURNAL OF TH

AMERICAN ACCOUNTING ASSOCIATIO

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2002–2003 Executive Committee of the Association are as follows:

President	G. Peter Wilson, Boston College
President-Elect	William L. Felix, Jr., The University of Arizona
Vice President	William E. (Bill) McCarthy, Michigan State University
Vice President–Education	Kevin D. Stocks, Brigham Young University
Vice President–Finance	Judy D. Rayburn, University of Minnesota
Vice President–International	Judy S. L. Tsui, The Hong Kong Polytechnic University
Vice President–Professional Relations	Samuel A. Vitkoski, BDO Seidman
Vice President–Publications	Lawrence A. Tomassini, The Ohio State University
Vice President–Research	Zoe-Vonna Palmrose, University of Southern California
Vice President–Sections and Regions	Joanna L. Ho, University of California, Irvine
Past President	Joel S. Demski, University of Florida

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive *Accounting Education News*, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues are based on the selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.*
(This does not include Section journals.)

Select one Association-wide journal	U.S. $85.00
Select two Association-wide journals	U.S. $95.00
All three Association-wide journals	U.S. $105.00

Associate membership for students residing in the United States and Canada:

Select one Association-wide journal	U.S. $20.00
Select two Association-wide journals	U.S. $30.00
All three Association-wide journals	U.S. $40.00

Electronic journal option available for an additional U.S. $20.00 to above-mentioned full-membership and associate-membership dues.

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Artificial Intelligence/Emerging Technologies; Auditing; Financial Accounting and Reporting; Gender Issues in Accounting; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Teaching and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to taruw@u.washington.edu or sent to Professor Terry Shevlin, Senior Editor, *The Accounting Review*, School of Business, Box 353200, University of Washington, Seattle, WA 98195-3200.

The submission fee of $75.00 for members of the AAA or $100.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm must be paid when article submitted. Books for review should be sent to Jane L. Reimers, College of Business, Florida State University, Tallahassee, FL 32306-1042. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review.*

The Accounting Review (ISSN 0001-4826) is published four times a year in January, April, July, and October by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rate of U.S. $175; membership subscription rate is $25 with discounts available; and emeritus subscription rate is $8.75. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to ***The Accounting Review,*** 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2003 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Terry Shevlin
Editors
Lawrence Brown
Patricia Dechow
S. Jane Kennedy
Marlys Gascho Lipe
Madhav Rajan

VOL. 78 APRIL 2003 No. 2

ASHIQ ALI, LEE-SEOK HWANG, and MARK A. TROMBLEY
Residual-Income-Based Valuation Predicts Future Stock Returns: Evidence on Mispricing vs. Risk Explanations 377
PETER F. CHEN and GUOCHANG ZHANG
Heterogeneous Investment Opportunities in Multiple-Segment Firms and the Incremental Value Relevance of Segment Accounting Data 397
ANDREW FERGUSON, JERE R. FRANCIS, and DONALD J. STOKES
The Effects of Firm-Wide and Office-Level Industry Expertise on Audit Pricing 429
BJORN N. JORGENSEN and MICHAEL T. KIRSCHENHEITER
Discretionary Risk Disclosures 449
KRISHNA R. KUMAR and GNANAKUMAR VISVANATHAN
The Information Content of the Deferred Tax Valuation Allowance 471
JOHN PHILLIPS, MORTON PINCUS, and SONJA OLHOFT REGO
Earnings Management: New Evidence Based on Deferred Tax Expense 491
STEPHEN G. RYAN and PAUL A. ZAROWIN
Why Has the Contemporaneous Linear Returns-Earnings Relation Declined? 523
KAREN L. SEDATOLE
The Effect of Measurement Alternatives on a Nonfinancial Quality Measure's Forward-Looking Properties 555
HUN-TONG TAN and KEN T. TROTMAN
Reviewers' Responses to Anticipated Stylization Attempts by Preparers of Audit Workpapers 581
Editorial Data 605
Editorial Policy and Style Information 606

Editorial Staff of *The Accounting Review*
2002–2003

SENIOR EDITOR
Terry Shevlin, University of Washington
EDITORS
Lawrence Brown, Georgia State University
Patricia Dechow, University of Michigan
S. Jane Kennedy, University of Washington
Marlys Gascho Lipe, University of Oklahoma
Madhav Rajan, Stanford University
BOOK REVIEW EDITOR
Jane L. Reimers, Florida State University

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland
COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Kathy Casper
Helen Griffith
Judy Cothern
Jan Kovarik (subcontracted)
OTHER AAA STAFF
Debora Gardner
Roland LaTulip
Deirdre Harris
Mary Cole
James Szelminski
Linda Watterworth
Myriam Chacon
Barbara Brady

Editorial Advisory and Review Board

Shannon W. Anderson, Rice University
Benjamin C. Ayers, University of Georgia
Frances L. Ayres, University of Oklahoma
William R. Baber, The George Washington University
Stephen P. Baginski, Indiana University
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Orie E. Barron, Pennsylvania State University
Eli Bartov, New York University
Anne L. Beatty, Pennsylvania State University
Robert J. Bloomfield, Cornell University
Robert M. Bowen, University of Washington
David C. Burgstahler, University of Washington
Brian J. Bushee, University of Pennsylvania
Robert M. Bushman, The University of North Carolina at Chapel Hill
C. Bryan Cloyd, University of Illinois at Urbana–Champaign
Mark L. DeFond, University of Southern California
Ilia D. Dichev, University of Michigan
Sunil Dutta, University of California, Berkeley
Merle M. Erickson, University of Chicago
Jere R. Francis, University of Missouri–Columbia
Paul Fischer, Pennsylvania State University
Jennifer Francis, Duke University
Jonathan C. Glover, Carnegie Mellon University
Wayne R. Guay, University of Pennsylvania
David A. Guenther, University of Colorado at Boulder
Mahendra Gupta, Washington University
D. Eric Hirst, The University of Texas at Austin
Christopher D. Ittner, University of Pennsylvania
James J. Jiambalvo, University of Washington
Deen Kemsley, Columbia University
William R. Kinney, Jr., The University of Texas at Austin
Lisa L. Koonce, The University of Texas at Austin
William N. Lanen, University of Michigan
Robert C. Lipe, University of Oklahoma
Joan L. Luft, Michigan State University
Russell J. Lundholm , University of Michigan
Joseph Magliolo III, Southern Methodist University
Laureen A. Maines, Indiana University
Steven R. Matsunaga, University of Oregon
Maureen F. McNichols, Stanford University
Donald V. Moser, University of Pittsburgh
James N. Myers, University of Illinois at Urbana–Champaign
Mark W. Nelson, Cornell University
Patricia C. O'Brien, University of Waterloo
Kathy R. Petroni, Michigan State University
Morton Pincus, The University of Iowa
Jamie H. Pratt, Indiana University
D. Shores, University of Washington
Richard G. Sloan, University of Michigan
Geoffrey B. Sprinkle, Indiana University
Theodore Sougiannis, University of Illinois at Urbana–Champaign
Hung-Tong Tan, Nanyang Technology University / Singapore
Richard M. Tubbs, The University of Iowa
Sandra Vera-Muñoz, University of Notre Dame
Robert E. Verrecchia, University of Pennsylvania
Beverly R. Walther, Northwestern University
Hong Xie, University of Illinois at Urbana–Champaign

PAST EDITORS

1926–1929	William A. Paton	1973–1975	Thomas F. Keller
1929–1943	Eric L. Kohler	1976–1978	Don T. DeCoster
1944–1947	A. C. Littleton	1978–1983	Stephen A. Zeff
1948–1949	Robert L. Dixon	1983–1987	Gary L. Sundem
1950–1959	Frank P. Smith	1987–1990	William R. Kinney, Jr.
1960–1962	Robert K. Mautz	1990–1994	A. Rashad Abdel-khalik
1963–1964	Lawrence L. Vance	1994–1997	Robert P. Magee
1965–1967	Wendell Trumbull	1997–2000	Gerald L. Salamon
1968–1970	Charles H. Griffin	2000–2002	Linda Smith Bamber
1971–1972	Eldon S. Hendriksen		

THE ACCOUNTING REVIEW
Vol. 78, No. 2
2003
pp. 377–396

Residual-Income-Based Valuation Predicts Future Stock Returns: Evidence on Mispricing vs. Risk Explanations

Ashiq Ali
The University of Arizona

Lee-Seok Hwang
Baruch College, The City University of New York and The Chinese University of Hong Kong (visiting)

Mark A. Trombley
The University of Arizona

ABSTRACT: Frankel and Lee (1998) show that the value-to-price ratio (V_f/P) predicts future abnormal returns for up to three years, where V_f is an estimate of fundamental value based on a residual income valuation framework operationalized using analyst earnings forecasts. In this study, we examine whether the V_f/P effect is due to market mispricing or omitted risk factors. We find that the V_f/P effect is partially concentrated around the future earnings announcements, consistent with the mispricing explanation. On using an extensive set of risk proxies, suggested by Gebhardt et al. (2001) and Gode and Mohanram (2001), we also find that V_f/P is significantly related to some risk proxies. However, after controlling for these risk factors, V_f/P continues to exhibit a significant positive association with future returns suggesting that these risk factors are not responsible for the V_f/P effect. Overall, the results seem consistent with the mispricing explanation for the V_f/P effect.

Keywords: *residual income valuation; mispricing; risk.*

Data Availability: *All data are obtained from publicly available sources.*

We thank Christi Gleason, Alister Hunt, Charles Lee, Lil Mills, Lynn Rees, Shiva Sivaramakrishnan, Lenny Soffer, Hong Xie, and participants in accounting research workshops at The University of Arizona, Texas A&M University, and the 2001 American Accounting Association Annual Meeting, and an anonymous referee for helpful comments. The analyst estimates data used in this study were provided by I/B/E/S under its program to encourage research on earnings expectations. Professors Ali and Trombley acknowledge financial support provided by The University of Arizona Eller Faculty Fellowships and by Ernst & Young. Earlier versions of the paper were titled: "Residual-Income-Based Valuation and Future Stock Returns: Tests of the Market Mispricing Hypothesis."

Editor's note: This paper was accepted by Daniel B. Thornton, *Ad Hoc* Editor.

Submitted May 2002
Accepted December 2002

I. INTRODUCTION

Frankel and Lee (1998) show that the value-to-price ratio (V_f/P) can predict abnormal returns for up to three years. They estimate fundamental value (V_f) by using the residual income model (Ohlson 1995), operationalized with analyst earnings forecasts. They argue that their study features a more complete and more robust valuation model than simple market-multiples, and that therefore their measure leads to relatively better cross-sectional predictions of future abnormal returns. They also show that their results are not attributable to certain risk proxies (book-to-market value of equity, firm size, and beta), and conclude that the V_f/P effect may reflect temporary mispricing by the market. However, they note the possibility of an alternative explanation: high V_f/P firms may still be riskier than low V_f/P firms in some other dimensions, causing V_f/P to be related to future returns. Beaver (2002), Kothari (2001), and Lo and Lys (2000) express a similar concern, and note that it remains to be seen whether market mispricing is the correct interpretation of the Frankel and Lee (1998) results. Our study provides evidence on the mispricing vs. risk explanations for the ability of V_f/P to predict abnormal returns.

We find that the V_f/P effect is partially concentrated around subsequent earnings announcements—evidence that is consistent with market mispricing and its subsequent correction when earnings information is released. We also examine whether the V_f/P effect is due to omitted risk factors for which V_f/P may be a proxy. Gebhardt et al. (2001) and Gode and Mohanram (2001) suggest several firm characteristics as proxies for risk. We find that V_f/P is significantly related to some of the risk factors, but that the direction of the relationships does not conclusively indicate whether high V_f/P firms are riskier. However, after controlling for all suggested risk factors, V_f/P continues to exhibit a significant positive association with future returns. Thus, although V_f/P is correlated with some risk factors, it does not appear to be driven by omitted risk factors. Overall, our evidence is consistent with the mispricing explanation rather than with the risk explanation for the V_f/P effect.

Several other influential studies in addition to Frankel and Lee (1998) underscore the importance of our evidence on the mispricing and risk explanations for the V_f/P effect. These studies have also documented the ability of residual-income-model-based fundamental value to predict future returns. Dechow et al. (1999) use historical earnings information to predict future returns and residual income, while Frankel and Lee (1998) use analysts' earnings forecasts to implement the residual income model. Unlike these two studies, which examine the ability of the residual income model to explain future returns cross-sectionally, Lee et al. (1999) and Lee and Swaminathan (1999) examine the time-series relation between fundamental value and price and fundamental value and returns, respectively, for several U.S. stock indices. Although all of these studies recognize the importance of understanding whether mispricing is a reason for the predictive ability of the fundamental-value-to-price ratio, none investigate this issue. Lee et al. (1999, 1694) argue in support of the mispricing explanation: "When intrinsic values are difficult to measure and/or when trading costs are significant, the process by which price adjusts to intrinsic value requires time, and price does not always perfectly reflect intrinsic value." They conclude by stating, "We leave the exact explanation for the predictive power of V/P (fundamental value to price) to future research" (Lee et al. 1999, 1737).

The remainder of the paper is organized as follows: Section II describes the Frankel and Lee (1998) V_f/P measure and documents V_f/P-based predictable abnormal returns. Section III describes empirical tests of the mispricing hypothesis and documents the results. Section IV presents tests of the risk explanation for the V_f/P effect. Section V compares the ability of alternate valuation measures in the literature, specifically those proposed by

Dechow et al. (1999) and Gode and Mohanram (2001), to predict future abnormal returns with the predictive ability of V_f/P. Section VI presents the summary and conclusions.

II. FUNDAMENTAL VALUATION AND ABNORMAL RETURN PREDICTABILITY

The V_f/P Measure

Motivated by recent theoretical papers on residual income valuation (e.g., Ohlson 1990, 1995), Frankel and Lee (1998) construct a measure of fundamental equity value. This approach models the value of equity as book value plus the present value of expected residual income, where residual income is the difference between reported income and the cost of equity capital multiplied by reported book value at the beginning of the period. The residual income model is derived from the discounted dividend valuation model and the assumption of clean-surplus accounting.[1] In practice, application of residual income valuation requires estimates of future income (or return on equity), cost of capital, and dividend payout rates, as well as an assumption about terminal value at the end of the forecasting horizon (Lee 1996). Frankel and Lee's (1998) primary measure of firm value uses the following model:

$$V_f = B_t + \frac{(FROE_t - r_e)}{(1 + r_e)} B_t + \frac{(FROE_{t+1} - r_e)}{(1 + r_e)^2} B_{t+1} + \frac{(FROE_{t+2} - r_e)}{(1 + r_e)^2 r_e} B_{t+2}$$

where $FROE_t$ is the forecasted return on equity in year t, B_t is the estimated book value of common stockholders' equity at the end of year t, and r_e is the estimated cost of capital. This formulation of the residual income model assumes that the forecasted residual income in year $t+2$ continues in perpetuity.

Frankel and Lee (1998) estimate $FROE_t$, $FROE_{t+1}$, and $FROE_{t+2}$ using one-year ($FY1$) and two-year ($FY2$) I/B/E/S consensus earnings estimates, five-year (long-term) I/B/E/S consensus growth rates (Ltg), actual book values (B_{t-2} and B_{t-1}), and estimated future book values (B_t, B_{t+1}):[2]

$$FROE_t = FY1/[(B_{t-1} + B_{t-2})/2]$$
$$FROE_{t+1} = FY2/[(B_t + B_{t-1})/2]$$
$$FROE_{t+2} = [FY2(1 + Ltg)]/[(B_{t+1} + B_t)/2].$$

Cost of capital (r_e) is estimated as an industry-specific rate using the three-factor industry risk premiums described in Fama and French (1997) plus an estimated annualized risk-free rate based on average 30-day Treasury-bill rates over our sample period.

---

[1] The clean-surplus assumption requires that the book-value of common equity at time t is equal to book value of common equity at time $t-1$ plus income in period t, adjusted by payments to or from stockholders during period t.

[2] Following Frankel and Lee (1998), where Ltg is not available, $FROE_{t+2}$ is set equal to $FROE_{t+1}$, and future book values are estimated using the clean surplus relation, $B_t = B_{t-1} + FY1 - DIV_t$, where dividends ($DIV_t$) are estimated based on the dividend payout ratio (or a proxy based on normalized earnings for loss firms) in the most recent year.

In computing $FROE_t$, Frankel and Lee (1998) use the average of B_{t-2} and B_{t-1}, rather than B_{t-1}, to reduce the chance of an extremely low denominator. However, since B_{t-1} is conceptually a more appropriate denominator to compute $FROE_t$, we repeat our analysis after making this as well as similar changes to $FROE_{t+1}$ and $FROE_{t+2}$. As expected, we delete about a half-percent more observations because of the extreme values of $FROE_t$, $FROE_{t+1}$, and $FROE_{t+2}$ (±1 is the cut-off value used in Frankel and Lee [1998] and in this study). However, the results remain very similar to those reported later in this paper.

V_f/P and Subsequent Abnormal Returns

Frankel and Lee (1998) form portfolios based on the ratio of V_f to market price (V_f/P) as of the end of June of each year during the period 1977 to 1992. They find that V_f/P predicts abnormal returns over one-, two-, and three-year holding periods (see Frankel and Lee 1998, Table 3, Panel D). Specifically, a portfolio constructed by taking a long position in firms in the highest quintile of V_f/P and a short position in firms in the lowest quintile of V_f/P produces returns of 3.1 percent, 15.2 percent, and 30.6 percent over one-, two-, and three-year holding periods, respectively.

We replicate the Frankel and Lee (1998) results using data from the period 1976 to 1997, which subsumes the period from which the Frankel and Lee data are collected. Other than the difference in data periods, our measurement, data sources, and portfolio formation approaches are consistent with those used by Frankel and Lee (1998). Panel A of Table 1 shows that a portfolio constructed by taking a long position in firms in the highest quintile of V_f/P, Q5, and a short position in firms in the lowest quintile of V_f/P, Q1, (hereafter, Q5-Q1 portfolio) produces returns of 5.7 percent, 15.5 percent, and 31.1 percent over one-, two-, and three-year holding periods (*Ret12*, *Ret24*, and *Ret36*), respectively. As in Frankel and Lee (1998), the V_f/P-based portfolio selection strategy results in statistically significant positive returns ($p < 0.01$) that increase over longer holding periods.[3]

Table 1 indicates that market capitalization (*ME*) is significantly greater for the high V_f/P quintile than for the low V_f/P quintile ($1,756 million vs. $834 million). To control for the effect of this size difference, we calculate size-adjusted returns for each V_f/P quintile, equal to raw returns minus the corresponding CRSP size-decile index returns. The Q5-Q1 portfolio produces size-adjusted returns of 4.7 percent, 12.5 percent, and 23.7 percent over one-, two-, and three-year holding periods (*SRet12*, *SRet24*, and *SRet36*), respectively. Thus, returns to the V_f/P-based portfolio selection strategy do not seem to result from size differences between the high V_f/P and low V_f/P portfolios.[4] Overall, the V_f/P effect for our sample is very consistent with that documented by Frankel and Lee (1998).[5]

III. TESTS OF THE MARKET MISPRICING HYPOTHESIS

If abnormal returns to the V_f/P strategy are due to omitted risk factors, then future abnormal returns should not be concentrated around the earnings announcement periods (Bernard and Thomas 1989, 1990; Bernard et al. 1997), because asset-pricing models do not predict significant shifts in expected returns over short windows. Thus, if we find evidence of concentration of returns around earnings announcements, this would suggest that the abnormal return predictability of V_f/P is due to mispricing and that the mispricing is corrected during subsequent earnings announcements because a relatively large amount of information reaches the market during earnings announcement periods.

Panel B of Table 1 reports that the Q5-Q1 portfolio returns during three-day windows of subsequent quarterly earnings announcements for one-, two-, and three-year horizons (*EARet12*, *EARet24*, *EARet36*) are 0.3 percent, 1.5 percent, and 2.4 percent, respectively,

---

[3] We compute portfolio returns for each sample year and use the time-series variation in the annual estimates to compute significance level for the mean portfolio return over the sample period. To correct for serial correlation in returns induced by overlapping holding periods for return horizons longer than one year, we use the Newey and West (1987) procedure. This procedure to compute the test statistics is used throughout the paper.

[4] B/P and *Beta* are also significantly different across the V_f/P quintiles. We control for these factors in a regression model, which is reported later.

[5] Herzberg (1998) also documents the robustness of the $V_{f/P}$ effect. He shows that the V_f/P effect is not sensitive to holding periods of different lengths and to two different measures of discount rates. He also shows that the effect occurs across a broad set of industries.

TABLE 1
Characteristics of V_f/P Quintile Portfolios Formed at the End of June Each Year over the 1976–1997 Period

Variable	Q1 (Low V_f/P)	Q2	Q3	Q4	Q5 (High V_f/P)	All Firms	Q5-Q1 Diff.
Panel A: Firm Characteristics and Returns by Quintile							
V_f/P	0.482	0.824	1.024	1.329	2.503	1.140	2.020***
B/P	0.668	0.644	0.721	0.737	0.711	0.661	0.042
ME	834	1,190	1,157	1,126	1,756	1,306	922***
Beta	1.044	1.008	0.997	1.044	1.149	1.031	0.104***
Ret12	0.137	0.156	0.164	0.172	0.194	0.165	0.057***
Ret24	0.257	0.313	0.322	0.342	0.412	0.324	0.155***
Ret36	0.411	0.486	0.493	0.550	0.723	0.518	0.311***
SRet12	−0.010	0.002	0.008	0.016	0.036	0.003	0.047***
SRet24	−0.031	0.002	0.008	0.031	0.093	0.007	0.125***
SRet36	−0.052	−0.010	0.001	0.058	0.185	0.021	0.237***
Panel B: Earnings Announcement Window Returns by Quintile							
EARet12	0.005	0.007	0.004	0.006	0.008	0.008	0.003
EARet24	0.011	0.015	0.009	0.017	0.027	0.020	0.015***
EARet36	0.018	0.022	0.017	0.030	0.043	0.033	0.024***
EASRet12	0.001	0.003	0.002	0.003	0.005	0.004	0.003
EASRet24	0.004	0.008	0.003	0.009	0.019	0.012	0.015**
EASRet36	0.007	0.012	0.008	0.017	0.032	0.021	0.024***
Obs.	4,017	4,031	4,032	4,031	4,022	20,133	

, * Statistical significance at the 5 percent and 1 percent levels, respectively, (two-tailed).

t-statistics are computed as the mean divided by the standard error of the annual values of the variables. The Newey and West (1987) procedure is used to adjust for serial correlation in the two- and three-year portfolio (size-adjusted) returns induced by overlapping holding periods.

For computing V_f, a firm-year observation is deleted from the sample if (1) the absolute value of future return-on-equity in years t, $t+1$, and $t+2$ is greater than or equal to 1, (2) dividend payout ratio is greater than or equal to 1, or (3) book value per share of common equity in years $t-1$ through $t+1$ is negative.

V_f/P is the analyst forecast-based fundamental value divided by stock price (P) as of the end of June of year t; *ME* is the market value of equity in millions of dollars at the end of June of year t; *B/P* is the book value per share for fiscal year that ends in year $t-1$ divided by price at the end of June of year t; *Beta* is the systematic risk estimated using monthly returns over a maximum of 36 months beginning July of year t; *Ret12*, *Ret24*, and *Ret36* are the one-year, two-year, and three-year buy-and-hold return, respectively, beginning July of year t; *SRet12*, *SRet24*, and *SRet36* are the size-adjusted one-year, two-year, and three-year buy-and-hold returns, respectively, beginning July of year t, defined as raw buy-and-hold return less the corresponding size-decile index return where size decile is based on size-decile cutoffs for all NYSE and AMEX firms; *EARet12*, *EARet24*, and *EARet36* (*EASRet12*, *EASRet24*, and *EASRet36*) are the raw (size-adjusted) one-year, two-year, and three-year total returns in the three-day periods (−1, 0, +1) around quarterly earnings announcements; and Obs. is the number of observations and applies to all variables except for earnings announcement return variables, which have fewer observations because of data availability.

with the last two numbers significant at better than the 1 percent level. Results are nearly identical when size-adjusted returns are used instead of raw returns. Abnormal returns that are significantly greater than zero around earnings announcement dates suggest that the V_f/P effect is more likely to be due to market mispricing than to omitted risk factors.[6]

The preceding portfolio analysis considers only firms in extreme quintiles of V_f/P. We employ regression analysis using firms in all five quintiles of V_f/P to evaluate the consistency of the relation between V_f/P and future returns. The regression framework also allows us to control for the effect of size on the concentration of returns around earnings announcements. Specifically, we estimate the following model:

$$EASRet = a_0 + a_1 RK(V_f/P) + a_2 Ln(ME) + e. \tag{1}$$

EASRet is size-adjusted return beginning one day prior to the announcement of earnings for a quarter and ending one day after the announcement. We calculate *EASRet* for each of the 12 quarters after the V_f/P measurement date. Thus, data permitting, 12 observations correspond to each V_f/P value. $RK(V_f/P)$ is the quintile rank of V_f/P scaled from 0 to 1, with rank assigned annually. We use a rank measure to reduce the effect of outliers and because of the possibility of nonlinearity in the relation between *EASRet* and V_f/P. *ME* is the market capitalization of the firm at the end of June of the year in which V_f/P is measured. *Ln(ME)* is included in Model (1) to protect the coefficient on $RK(V_f/P)$ from omitted variable bias. As noted in the discussion of Table 1, V_f/P is positively related to firm size, and small firms tend to have positive abnormal returns on earnings announcement dates (Chari et al. 1988; Ball and Kothari 1991). Thus, if firm size is not controlled for, the coefficient on $RK(V_f/P)$ in Equation (1) could be biased.

We estimate Equation (1) annually and report in Table 2 the cross-temporal means of the regression coefficients and the Newey and West (1987) adjusted t-statistics of the means. The coefficient on $RK(V_f/P)$ is positive and significant, 0.178 (t = 2.887).[7] As an additional test of the significance of the regression coefficient on $RK(V_f/P)$, we conduct the following simulation. For each firm quarter, we randomly select a three-day period from the non-earnings-announcement days starting two days after the announcement of previous quarter earnings and ending two days before the announcement of current quarter earnings. Using this random three-day period in place of the actual earnings announcement period, we estimate Equation (1) and obtain a value for a_1, the slope coefficient of $RK(V_f/P)$. We repeat this process 999 times. We find that the simulated values of the slope coefficient on $RK(V_f/P)$ exceed the actual slope coefficient on $RK(V_f/P)$, 0.178, only once (see Table 2), suggesting a significance level for a_1 of 0.002 (equal to $(1 + 1)/(999 + 1)$; see Noreen

---

[6] We do not conduct our analysis with a wider event window (longer than three days), as did Skinner and Sloan (2002). They predict that the entire return differential between growth stocks and other stocks will be realized during periods in which negative earnings surprises are reported. Given that negative earnings news are generally predisclosed before the earnings announcement date, it is important for Skinner and Sloan (2002) to consider the pre-announcement period returns. For the V_f/P effect, we find (in an analysis similar to Skinner and Sloan [2002], not reported in the paper) that the stock price reactions to negative earnings surprises are not asymmetrically large, so there does not seem to be any clear benefit of using a longer event window, especially given that the longer event window decreases the usefulness of the short-window as a means of controlling expected returns.

[7] This coefficient indicates that quarterly earnings announcement return to a hedge portfolio long in the high V_f/P firms and short in the low V_f/P firms is 0.178 percent. Over 12 quarters, this return is equivalent to 2.136 percent (0.178 × 12), which is very consistent with the 36-month hedge portfolio return of 2.4 percent, reported in Panel B of Table 2.

TABLE 2
Regressions of Full Quarter Stock Returns and Three-Day Earnings Announcement Period Stock Returns on V_f/P and Size

$$EASRet = a_0 + a_1 RK(V_f/P) + a_2 Ln(ME) + e \tag{1}$$

$$FullSRet = b_0 + b_1 RK(V_f/P) + b_2 Ln(ME) + e \tag{2}$$

	Dependent Variable	
	EASRet	*FullSRet*
Intercept	0.408*** (3.615)	0.445 (0.659)
$RK(V_f/P)$	0.178*** (2.887)	1.321*** (3.908)
$Ln(ME)$	−0.059*** (−3.361)	−0.162* (−1.674)
Adj. R^2	0.052%	0.194%
Obs.	193,911	193,911
Years	1976–1997	1976–1997
Test: a_1 − (5% of b_1)		0.112** (2.470)
		Frequency
Results based on 999 simulated regressions:		
$a_1 \geq 0.178$		1
Results based on 1000 simulated regressions:		
t-statistics for a_1 − 5% of b_1:		
Significant at better than the 10% level		55
Significant at better than the 5% level		36
Significant at better than the 1% level		14

---

*, **, *** Statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively, (two-tailed).

Regressions are estimated annually and means of the annual estimates are reported. t-statistics (in parentheses below estimated coefficients) are computed as the mean divided by the standard error of the annual estimates.

Simulated a_1 is obtained as follows: For each firm-quarter, we randomly select a three-day period from the non-earnings announcement days starting two days after the previous quarter earnings announcement date and ending two days before the current quarter's earnings announcement date. Using this random three-day period in place of the actual earnings announcement period, we estimate Equation (1) and obtain a value for a_1, the slope coefficient of $RK(V_f/P)$.

V_f/P is analyst forecast-based fundamental value divided by stock price as of the end of June in year t; *FullSRet* is size-adjusted return (in percentage) beginning two days after earnings announcement date of the previous quarter and ending one day after earnings announcement date of the current quarter; and *EASRet* is size-adjusted return (in percentage) beginning one day prior to the earnings announcement date and ending one day after earnings announcement date of the current quarter, where size-adjusted return is defined as raw buy-and-hold return less NYSE/AMEX size-decile return. A maximum of 12 quarterly *FullSRet* and 12 quarterly *EASRet* values correspond to each V_f/P value; $RK(V_f/P)$ is the quintile rank measure scaled from 0 to 1, in ascending order of V_f/P each year; and *ME* is the market capitalization of the firm at the end of June of year t.

1989, 44). These results are consistent with the presence of a significant level of V_f/P-related abnormal returns around earnings announcements.

Having established that the V_f/P-related abnormal returns around earnings announcement dates are significantly greater than zero, we now focus on how the magnitude of these abnormal returns compare with the magnitude of abnormal returns during the non-earnings-announcement period. For this purpose, we estimate the following model:

$$FullSRet = b_0 + b_1RK(V_f/P) + b_2Ln(ME) + e, \tag{2}$$

where *FullSRet* is size-adjusted return cumulated over the period beginning two days after the announcement of earnings for the previous quarter and ending one day after the announcement of earnings for the current quarter. We calculate *FullSRet* for each of the 12 quarters after the V_f/P measurement date. Thus, data permitting, 12 observations correspond to each V_f/P value. We include $Ln(ME)$ in Model (2) to make the coefficient on $RK(V_f/P)$ comparable across Models (1) and (2).

The coefficient a_1 in Equation (1) represents three-day earnings announcement date abnormal returns related to the V_f/P effect, and b_1 in Equation (2) represents three-month abnormal returns related to the V_f/P effect. Since *FullSRet* is measured over an average of 63 days and *EASRet* is measured over 3 days (approximately 5 percent of 63), if a_1 (the slope coefficient on $RK(V_f/P)$ in Equation (1)) is significantly greater than 5 percent of b_1 (the slope coefficient on $RK(V_f/P)$ in Equation (2)), then this would indicate that the V_f/P effect is concentrated in the earnings announcement periods. Note that Equations (1) and (2) are estimated using the same number of observations because we use only those observations for which data are available for both *EASRet* and *FullSRet*. This restriction makes the comparison of the coefficients across Equations (1) and (2) more meaningful.

Table 2 reports the regression estimates of Equation (2). The coefficient b_1 is positive and significant, 1.321 (t = 3.908), confirming the existence of the V_f/P effect. The results also indicate that the coefficient a_1 in the *EASRet* regression, 0.178, is significantly greater than 5 percent of b_1; the difference, a_1 − 5 percent of b_1, is 0.112, with a t-statistic of 2.470.[8] This result suggests that the returns associated with a V_f/P-based trading strategy are partially concentrated around earnings announcements and is consistent with the mispricing explanation.[9]

In the preceding test, we implicitly assume that the slope coefficient for any three-day period is expected to be 5 percent of the slope coefficient for the full quarter. If this assumption is not valid, the test statistic may reject the null hypothesis of (a_1 − 5 percent of b_1) = 0 too often and the reported t-statistic may be overstated. To validate this assumption, we do the following. Using the procedure described earlier, we obtain 1,000 simulated values of a_1 and examine how often the value of (a_1 − 5 percent of b_1) is rejected at particular levels of significance. We find that the test statistics are significant at better than the 10 percent (5 percent, 1 percent) level in 5.5 percent (3.6 percent, 1.4 percent) of the 1,000 trials. To be valid, the probability of falsely rejecting the null should be less than or equal to the rejection level of the test (Noreen 1989, 50), which it is for the 10 percent and 5 percent levels (though not for the 1 percent level). Overall, the simulation results suggest that the test statistics do not reject the null hypothesis of (a_1 − 5 percent of b_1) = 0 too often, and the t-statistic for (a_1 − 5 percent of b_1) reported in Table 2 is not likely to be overstated.

---

[8] Using the continuous measure of V_f/P instead of the rank measure, the estimates of a_1 and b_1 are 0.095 (t = 2.515) and 0.706 (t = 3.417), respectively, and the test a_1 − 5 percent of b_1 produces a test statistic of 2.509. These results are similar to those based on the rank measure.

[9] The returns associated with the V_f/P strategy around earnings announcements are 13.5 percent (0.178/1.321) of the total returns for announcement and non-announcement periods. For the book-to-market strategy, based on La Porta et al. (1997, Table I), the corresponding concentration ratio is 25.4 percent. One reason for the lower concentration for the V_f/P strategy is that the requirement for analysts' forecasts in computing V_f/P causes our sample to reduce by 58.6 percent, generally in favor of larger firms. For these firms, news gets incorporated in stock price more continuously rather than being heavily concentrated on earnings announcement days (Freeman 1987). The concentration ratio for the book-to-market strategy for large firms (size greater than NYSE median) is 13.2 percent (La Porta et al. 1997, Table III), which is comparable to our results.

IV. TESTS OF THE RISK HYPOTHESIS

An alternative explanation for the ability of V_f/P to predict abnormal returns is the incorrect measurement or omission of risk factors that affect expected returns (Ball et al. 1995; Kothari 2001). Given that the V_f/P measure is calculated using market price, and since risk is reflected in market price, V_f/P might simply proxy for expected returns (Ball 1978; Keim 1988); the numerator, V_f, may function as a cross-sectional scaling factor (Berk 1995; Biddle and Hunt 1997). To the extent that abnormal returns computed by researchers may not have properly considered all risk factors in estimating expected returns, V_f/P may exhibit a relation with abnormal returns due to the mismeasurement or omission of risk factors. One way to alleviate this concern is to regress estimated abnormal returns against V_f/P after controlling for the various risk proxies that have been identified in the literature. Frankel and Lee (1998) control for risk related to beta, firm size, and the book-to-market ratio. However, they recognize the possibility that high V_f/P firms may still be riskier than low V_f/P firms in some other dimensions, which could cause V_f/P to be related to future returns. Thus, we control for several additional firm characteristics that have been considered in recent research papers as proxies for risk (Gebhardt et al. 2001; Gode and Mohanram 2001).

In addition to price in the denominator of the V_f/P driving the relation with estimated future abnormal returns, error in the computation of V_f may also lead to a relation with estimated future abnormal returns. The computation of V_f requires an estimate of expected return to discount future residual income. To the extent that the estimated expected return has negative (positive) error, it overstates (understates) the V_f measure and simultaneously overstates (understates) the estimate of abnormal returns, leading to a spurious association between V_f and estimated future abnormal returns (Lo and Lys 2000). This problem may not be as severe in our study because we do not use the same measure of expected return to calculate abnormal return (in computing *SRet36*, we use NYSE/AMEX size-decile index returns as expected returns) and to calculate V_f (we use the Fama and French [1997] three-factor approach to compute r_e, expected return to discount future residual income). In any case, we address this concern in our analysis.

To explore the risk explanation for the V_f/P effect, we first examine whether V_f/P is in fact related to firm characteristics that have been shown in the literature to be risk proxies. Specifically, we estimate the following model:

$$\begin{aligned} V_f/P = {} & a_0 + a_1 Beta + a_2 Ivolatility + a_3 D/M + a_4 Ln(ME) + a_5 Analysts \\ & + a_6 Altman\text{'s } Z + a_7 Disp + a_8 StdROA + a_9 ImpliedCC + a_{10} B/P \\ & + a_{11} Ltg + a_{12} r_e + e. \end{aligned} \tag{3}$$

The explanatory variables we include in the model are primarily motivated by Gebhardt et al. (2001) and Gode and Mohanram (2001). *Beta* and *Ivolatility* represent the systematic and non-systematic components of stock price variability, respectively. *Ivolatility* is included in this analysis under the assumption that the market prices idiosyncratic risk, at least to some extent (Merton 1987). Debt-equity ratio (*D/M*) captures the effect of greater leverage on firm risk (Fama and French 1992). *Altman's Z* captures the risk of financial distress (Altman 1968). Size (*ME*) is associated with differences in information environment that can lead to higher risk for small firms than for large firms. Similarly, the number of analysts following the firm (*Analysts*) is associated with the amount and quality of information (Brennan et al. 1993). Less predictable earnings and greater earnings variability result in greater risk of unfavorable earnings news; these factors are proxied by *Disp* (the dispersion of analysts' earnings forecasts), and *StdROA* (the standard deviation of past return-on-asset rates). We also include *ImpliedCC*, the average industry cost of capital that Gebhardt et al. (2001, Table 2) report for their sample period, 1979 to 1995, which is very similar to our 1976 to 1997 sample period. To the extent that the market compensates stockholders for risks not captured by the firm characteristics we have included in Equation (3), *ImpliedCC* may capture those omitted risk factors. Finally, we include three risk measures that are also used in the construction of V_f/P: B/P, *Ltg*, and r_e. B/P is the book-to-price ratio, *Ltg* is long-term growth forecasts from I/B/E/S, and r_e is estimated expected return. Although theoretical arguments do not suggest B/P and *Ltg* as risk factors, the fact that they have been found to be associated with future returns has led some to argue that they may be proxies for omitted risk factors (see, e.g., Fama and French 1992).

We report in Table 3 correlations among all the variables in Equation (3). Positive and significant correlations (both Pearson and Spearman) of V_f/P with *Beta*, *ImpliedCC*, and B/P suggest that high V_f/P stocks are riskier. However, negative and significant correlations (both Pearson and Spearman) of V_f/P with *Disp* and r_e and a positive and significant correlation of V_f/P with *Ltg* suggest that high V_f/P stocks are less risky. Thus, the univariate correlation results are inconclusive on whether omitted risk factors are responsible for high V_f/P firms exhibiting high future returns.

In Table 4, we report regression results for models based on Equation (3). The model in column 1 has B/P, *Ltg*, and r_e, the factors used in constructing V_f/P, as explanatory variables. The model in column 2 has the remaining risk factors as explanatory variables, and the model in column 3 includes all of the risk factors. The signs and significance levels of the coefficients are very similar across the three models, hence we discuss the results of only the full model. The coefficients on *Beta*, D/M, *Altman's Z*, *ImpliedCC*, and B/P are positive and significant (t-statistics ranging in magnitude from 2.486 to 8.032), suggesting that firms with high V_f/P are riskier and should therefore be associated with higher future returns. However, the coefficients on *Analysts* and *Ltg* are positive and significant (t = 3.703 and 16.930, respectively), and the coefficients on *Disp* and r_e are negative and significant (t = −2.485 and −10.371, respectively), suggesting that high V_f/P firms bear lower risk and should therefore be associated with lower future returns. Note that the high level of significance on the coefficients on B/P, *Ltg*, and r_e (t = 7.265, 16.930, and −10.371, respectively) is probably because these variables are used in the construction of V_f/P. As with the correlation results in Table 3, the regression results do not conclusively suggest whether high V_f/P firms are riskier. Therefore, we cannot conclude whether omitted risk factors are driving the V_f/P effect.

Next we examine whether V_f/P is related to future returns after controlling for the risk factors considered in Equation (3). If the relation remains significant after controlling for these risk factors, then this would indicate that the V_f/P effect is not driven by the omitted risk factors. Specifically, we estimate the following model:

$$\begin{aligned} SRet36 = {} & a_0 + a_1 V_f/P + a_2 Beta + a_3 Ivolatility + a_4 D/M + a_5 Ln(ME) \\ & + a_6 Analysts + a_7 Altman's\ Z + a_8 Disp + a_9 StdROA + a_{10} ImpliedCC \\ & + a_{11} B/P + a_{12} Ltg + a_{13} r_e + e \qquad (4) \end{aligned}$$

The inclusion of r_e in Equation (4) also helps address the Lo and Lys (2000) concern that an error in estimating r_e may induce a spurious correlation between V_f and future abnormal returns.

TABLE 3

Pearson and Spearman Rank Correlations among V_f/P and Risk Proxies

	V_f/P	*Beta*	*Ivolatility*	*D/M*	*Ln(ME)*	*Analysts*	*Altman's Z*	*Disp*	*StdROA*	*ImpliedCC*	*Ltg*	*B/P*	r_e
V_f/P		0.054***	0.019	0.009	0.038***	0.091***	0.092***	−0.058***	0.005	0.058***	0.281***	0.039**	−0.250***
Beta	0.031**		0.202***	−0.030**	−0.080***	−0.040**	−0.006	0.057***	0.169***	0.051***	0.127***	−0.048***	−0.002
Ivolatility	0.019	0.253***		0.062***	−0.591***	−0.402***	0.012	0.237***	0.406***	0.058***	0.246***	0.075***	−0.066***
D/M	0.064***	−0.058***	−0.060**		−0.120***	−0.065***	−0.193***	0.206***	−0.070***	−0.006	−0.036**	0.559***	0.134***
Ln(ME)	−0.052***	−0.046	−0.627***	−0.052***		0.795***	−0.157***	−0.131***	−0.224***	0.008	−0.175***	−0.222***	−0.075***
Analysts	−0.023	0.005	−0.467***	−0.026	0.806***		−0.176***	−0.048***	−0.154***	0.013	−0.127***	−0.113***	−0.164***
Altman's Z	0.016	0.078**	−0.368***	−0.038***	0.852***	0.787***		−0.094***	−0.053***	0.071***	0.005	−0.065***	0.124***
Disp	−0.134***	0.122***	0.329***	0.252***	−0.196***	−0.063***	−0.084***		0.222***	0.014	0.153***	0.225***	0.040**
StdROA	−0.017	0.201***	0.460***	−0.142***	−0.288***	−0.210***	−0.144***	0.383***		0.021***	0.150***	−0.023**	−0.064**
ImpliedCC	0.048***	0.047***	0.026**	−0.012	0.020**	0.026***	0.052***	−0.009	0.018***		0.027***	0.068***	0.083***
Ltg	0.224***	0.221***	0.349***	−0.217***	−0.218***	−0.152***	−0.095***	0.005	0.194***	0.021**		−0.041*	−0.001
B/P	0.067***	−0.104***	0.035	0.548***	−0.275***	−0.165***	−0.222***	0.369***	−0.017	0.046***	−0.270***		0.142***
r_e	−0.248***	0.032**	−0.043**	0.146***	−0.045***	−0.072***	−0.071***	0.064***	−0.004	0.139***	−0.023	0.145***	

*, **, *** Statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively, (two-tailed).

Pearson (Spearman) correlation coefficients are above (below) the diagonal. Correlation coefficients are calculated each year from 1976 to 1997 and means of the annual correlation coefficients are reported. Significance levels are computed using the mean and the standard error of the annual coefficients.

V_f/P is analyst forecast-based fundamental value divided by stock price as of the end of June in year t; *Beta* is systematic risk estimated using monthly returns over a maximum of 36 months ending in July of year t; *Ivolatility* is the standard deviation of residuals from a market model regression estimated using daily returns, over a one-year period ending in June of year t; *D/M* is the book value of long-term debt in year $t-1$ divided by market value of equity at the end of June of year t; *ME* is market value of equity in millions at the end of June of year t; *Analysts* is the number of analysts' estimates included in the I/B/E/S database in May of year t; *Altman's Z* is the score from Altman (1968) discriminant model: $Z = 0.012 \times$ (working capital/total assets) $+ 0.014 \times$ (retained earnings/total assets) $+ 0.033 \times$ (earnings before interest and taxes/total assets) $+ 0.006 \times$ (market value of equity/book value total liabilities) $+ 0.999 \times$ (sales/total assets), with all the variables from year $t-1$; *Disp* is dispersion in one-year-ahead analyst consensus earnings forecasts divided by absolute value of one-year-ahead analyst consensus earnings forecasts in May of year t; *StdROA* is standard deviation of return-on-assets in prior five years, $t-5$ to $t-1$; *ImpliedCC* is the industry implied cost-of-capital (Gebhardt et al. 2001, Table 2); *Ltg* is long-term growth in earnings estimate in percentage forecasted in May of year t; *B/P* is book value in year $t-1$ divided by market value of equity at the end of June of year t; and r_e is the industry cost of capital based on the Fama and French (1997) three-factor model, and is used in computing the V_f measure. The smallest (largest) 0.5% of the observations of each variable is set to the 0.005 (0.995) fractile value of the variable, except for *ImpliedCC* and r_e.

TABLE 4
Regression of V_f/P on Risk Proxies

$$V_f/P = a_0 + a_1 Beta + a_2 Ivolatility + a_3 D/M + a_4 Ln(ME) + a_5 Analysts + a_6 Altman's\ Z + a_7 Disp + a_8 StdROA + a_9 ImpliedCC + a_{10} B/P + a_{11} Ltg + a_{12} r_e + e \quad (3)$$

	Coefficient (t-statistic)		
	(1)	(2)	(3)
Intercept	2.174***	0.722***	1.675***
	(9.572)	(7.571)	(11.756)
Beta		0.040***	0.031***
		(3.033)	(3.566)
Ivolatility		7.537	−0.988
		(1.451)	(−0.340)
D/M		0.089**	0.103***
		(2.532)	(3.007)
Ln(ME)		−0.020	0.019
		(−1.605)	(1.341)
Analysts		0.020***	0.009***
		(5.857)	(3.703)
Altman's Z		0.136***	0.180***
		(6.387)	(8.032)
Disp		−0.189*	−0.452**
		(−1.837)	(−2.485)
StdROA		1.169*	1.179
		(1.786)	(1.532)
ImpliedCC		0.018	0.026**
		(1.548)	(2.486)
B/P	0.142***		0.171***
	(4.442)		(7.265)
Ltg	0.013***		0.015***
	(14.970)		(16.930)
r_e	−0.103***		−0.108***
	(−8.032)		(−10.371)
Adj. R^2	0.160	0.056	0.220
Obs.	18,199	18,199	18,199
Years	1976–1997	1976–1997	1976–1997

---

*, **, *** Statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively, (two-tailed).
Regression coefficients are estimated annually and means of the annual estimates are reported. t-statistics are computed as mean divided by standard error of the annual estimates.

V_f/P is analyst forecast-based fundamental value divided by stock price as of the end of June in year *t*; *Beta* is systematic risk estimated using monthly returns over a maximum of 36 months ending in July of year *t*; *Ivolatility* is the standard deviation of residuals from a market model regression estimated using daily returns, over a one-year period ending in June of year *t*; *D/M* is the book value of long-term debt in year $t-1$ divided by market value of equity at the end of June of year *t*; *ME* is market value of equity in millions at the end of June of year *t*; *Analysts* is the number of analysts' estimates included in the I/B/E/S database in May of year *t*; *Altman's Z* is a score from Altman (1968) discriminant model: Z = 0.012 × (working capital/total assets) + 0.014 × (retained earnings/total assets) + 0.033 × (earnings before interest and taxes/total assets) + 0.006 × (market value of equity/book value total liabilities) + 0.999 × (sales/total assets), with all the variables from year $t-1$; *Disp* is dispersion in one-year-ahead analyst consensus earnings forecasts divided by absolute value of one-year-ahead analyst consensus earnings forecasts in May of year *t*; *StdROA* is standard deviation of return-on-assets in prior five years, $t-5$ to $t-1$; *ImpliedCC* is the industry implied cost-of-capital (Gebhardt et al. 2001, Table 2); *Ltg* is long-term growth in earnings estimate in percentage forecasted in May of year *t*; *B/P* is book value in year $t-1$ divided by market value of equity at the end of June of year *t*; and r_e is the industry cost of capital based on the Fama and French (1997) three-factor model, and is used in computing the V_f measure. The smallest (largest) 0.5% of the observations of each variable are set to the 0.005 (0.995) fractile values of the variable, except for *ImpliedCC* and r_e.

Table 5 reports regression results for models based on Equation (4). The model in column 1 has only V_f/P as the explanatory variable. The slope coefficient is positive and significant (t = 3.155), confirming the presence of the V_f/P effect. The model in column 2 has B/P, *Ltg*, and r_e as explanatory variables. The coefficients on B/P and r_e are significant (t = 4.761 and −2.655). Thus, it is possible that the ability of V_f/P to predict returns may be driven by the ability of the factors used to construct V_f/P to predict returns. To examine this issue we add V_fP as an explanatory variable in the model in column 3. The

TABLE 5
Regression of Future Returns on V_f/P and Risk Proxies

$$\begin{aligned} SRet36 = {} & a_0 + a_1 V_f/P + a_2 Beta + a_3 Ivolatility + a_4 D/M + a_5 Ln(ME) + a_6 Analysts \\ & + a_7 Altman's\ Z + a_8 Disp + a_9 StdROA + a_{10} ImpliedCC + a_{11} B/P \\ & + a_{12} Ltg + a_{13} r_e + e \qquad (4) \end{aligned}$$

	Coefficient (t-statistic)			
	(1)	(2)	(3)	(4)
Intercept	−0.098** (−2.228)	0.168** (2.044)	−0.041 (−0.328)	0.265 (1.312)
V_f/P	0.120*** (3.155)		0.117*** (2.673)	0.108** (2.301)
Beta				0.072** (2.079)
Ivolatility				−6.671 (−1.606)
D/M				−0.014 (−0.636)
Ln(ME)				−0.025 (−1.429)
Analysts ($\times 10^{-2}$)				−0.342 (−1.503)
Altman's Z				0.006 (0.143)
Disp				0.079 (0.550)
StdROA				0.308 (0.523)
ImpliedCC				0.001 (0.113)
B/P		0.120*** (4.761)	0.108*** (4.556)	0.094*** (3.571)
Ltg ($\times 10^{-2}$)		−0.012 (−0.104)	−0.153 (−1.139)	−0.157 (−1.432)
r_e		−0.017*** (−2.655)	−0.007 (−0.878)	−0.012 (−1.255)
Adj. R^2	0.014	0.019	0.035	0.080
Obs.	18,199	18,199	18,199	18,199
Years	1976–1997	1976–1997	1976–1997	1976–1997

(*continued on next page*)

TABLE 5 (continued)

, * Statistical significance at the 5 percent and 1 percent levels, respectively, (two-tailed).

Regression coefficients are estimated annually and means of the annual estimates are reported. t-statistic is computed as mean divided by standard error of the annual estimates. The Newey and West (1987) procedure is used to adjust for serial correlation induced by three-year abnormal returns with overlapping periods.

SRet36 is the size-adjusted three-year buy-and-hold return beginning in July of year *t*, defined as raw buy-and-hold return less the corresponding NYSE-AMEX size-decile index return; V_f/P is analyst forecast-based fundamental value divided by stock price as of the end of June in year *t*; *Beta* is systematic risk estimated using monthly returns over a maximum of 36 months ending in July of year *t*; *Ivolatility* is the standard deviation of residuals from a market model regression estimated using daily returns, over a one-year period ending in June of year *t*; *D/M* is the book value of long-term debt in year $t-1$ divided by market value of equity at the end of June of year *t*; *ME* is market value of equity in millions at the end of June of year *t*; *Analysts* is the number of analysts' estimates included in the I/B/E/S database in May of year *t*; *Altman's Z* is a score from Altman (1968) discriminant model: $Z = 0.012 \times$ (working capital/total assets) $+ 0.014 \times$ (retained earnings/total assets) $+ 0.033 \times$ (earnings before interest and taxes/total assets) $+ 0.006 \times$ (market value of equity/book value total liabilities) $+ 0.999 \times$ (sales/total assets), with all the variables as of year $t-1$; *Disp* is dispersion in one-year-ahead analyst consensus earnings forecasts divided by absolute value of one-year-ahead analyst consensus earnings forecasts in May of year *t*; *StdROA* is standard deviation of return-on-assets in prior five years, $t-5$ to $t-1$; *Ltg* is long-term growth in earnings estimate in percentage forecasted in May of year *t*; *B/P* is book value in year $t-1$ divided by market value of equity at the end of June of year *t*; *ImpliedCC* is the industry implied cost-of-capital (Gebhardt et al. 2001, Table 2); and r_e is the industry cost of capital based on the Fama and French (1997) three-factor model, and is used in computing the V_f measure. The smallest (largest) 0.5% of the observations of each variable is set to the 0.005 (0.995) fractile value of the variable, except for *ImpliedCC* and r_e.

coefficient on V_f/P is significant (t = 2.673), suggesting that the ability of B/P and r_e to forecast future returns is not responsible for the V_f/P effect. This result in not surprising given that these variables explain only a small variation in V_f/P (adjusted R^2 is 16.7 percent in column 1 of Table 4). The results in column 3 of Table 5 also suggest that the V_f/P effect is not driven by a spurious association resulting from the measurement error in the discount factor r_e, as conjectured by Lo and Lys (2000). If that were the case, then the presence of r_e would have caused an insignificant coefficient on V_f/P. Finally, column 4 reports the results of the full model (Equation (4)). The coefficient on V_f/P remains positive and significant (t = 2.301), suggesting that omission of risk factors is not a likely explanation for the V_f/P effect.

In column 4, other than V_f/P, only two variables—*Beta* (t = 2.079) and B/P (t = 3.571)—have significant coefficients. The lack of significance of the other risk proxies may be caused by more than one variable representing the same risk characteristic (see the correlations in Table 3). The difference in the adjusted R^2 for the models in columns 1 and 4 (0.014 and 0.080, respectively) suggests that many of the risk variables, though not incrementally significant, may still have contributed nontrivially in explaining the variation in *SRet36*.

V. ALTERNATE VALUATION MEASURES

There are two alternative valuation measures in the literature that are conceptually similar to V_f/P—those suggested by Dechow et al. (1999) and Gode and Mohanram (2001). In this section, we compare the ability of these alternative measures to predict future abnormal returns with the predictive ability of V_f/P.[10]

---

[10] Modeling a fundamental valuation measure (*V*) and price (*P*) as a cointegrated system, Lee et al. (1999) argue that one indication of a superior estimate of *V* is that the associated *V/P* ratio will have a higher predictive power for future returns.

Dechow et al. (1999), like Frankel and Lee (1998), use the Ohlson (1995) residual income model, but operationalize it using historical data rather than analysts' forecasts. Specifically, they measure fundamental value, V_{dhs}, as follows:

$$V_{dhs} = B_t + \frac{(E_t - rB_{t-1})\omega_t}{1 + r - \omega_t}$$

where B_t is book value of common equity at time t, E_t is earnings at time t, r is the cost of equity capital, assumed to be 12 percent, and ω_t is a persistence factor for year t, obtained from the following pooled cross-sectional time-series regression using all firm-years with available data in all prior years up to year $t-1$:

$$(E_{i,t-1} - rB_{i,t-2}) = \alpha_t + \omega_t(E_{i,t-2} - rB_{i,t-3}) + e_{it-1}.$$

The Gode and Mohanram (2001) valuation measure is based on a model by Ohlson and Juettner-Nauroth (2000) and is operationalized as follows:

$$V_{gm} = \frac{eps_1}{r_e} + \frac{eps_2 - eps_1}{r_e(r_e - g_p)}$$

This valuation measure relies on two periods of earnings forecasts (eps_1 and eps_2), an expected perpetual growth rate beyond year two (g_p), and a discount rate (r_e). Following Gode and Mohanram (2001), we use I/B/E/S consensus one- and two-year ahead forecasts to measure eps_1 and eps_2 and set g_p = 4 percent. For r_e, we use the same discount rate measure as in V_f. Unlike V_f, V_{gm} neither uses book value as part of the calculation nor requires specific estimation of a dividend payout rate. The salient features of the three valuation measures, V_f, V_{dhs}, and V_{gm} are summarized in Table 6.

We compare the predictive ability of each of the three valuation measures by regressing them on future returns, *SRet36*. The first three columns of Table 7 report that V_f/P, V_{dhs}/P, and V_{gm}/P have positive and significant coefficients, 0.112 (t = 2.853), 0.115 (t = 2.915), and 0.021 (t = 2.030), respectively. This suggests that all three valuation measures have the ability to predict future returns. Although mean adjusted R^2 is greater for the V_f/P measure (0.014) than for V_{dhs}/P and V_{gm}/P measures (0.009 and 0.009), the differences are not statistically significant, based on the standard error of the annual differences. The lack of significance may be due to the lack of power of this test, and so we use the following alternate approaches to compare the predictive ability of the valuation measures.

Both V_f/P and V_{dhs}/P can be expressed as $V/P = B/P + PV(RI)/P$; where $PV(RI)$ is the present value of expected residual income. The $PV(RI)$ term is measured differently for V_f/P and V_{dhs}/P, while the B/P term is the same. In column 4 of Table 7, we report the results of regressing *SRet36* against V_f/P and B/P, and in column 5 the results of regressing *SRet36* against V_{dhs}/P and B/P. In these models, the predictive power of V_f/P and V_{dhs}/P comes from the respective $PV(RI)$ component. The results in column 4 shows that both V_f/P and B/P have significant coefficients (t = 2.574 and 3.990, respectively), suggesting that the $PV(RI)$ component contributes significantly in predicting future returns. The results in column 5 shows that the coefficient on B/P is significant (t = 4.384), but the coefficient on V_{dhs}/P is not (t = 1.600), suggesting that the $PV(RI)$ component does not contribute

TABLE 6
Salient Features of the Alternative Measures of Fundamental Value

Valuation Measure	Book Value of Equity Used	Forecasts	Discount Rate	Terminal Value
V_f (Frankel and Lee 1998)	Yes	Analysts' earnings forecasts	Fama and French (1997) three-factor, by industry	Residual income of $t+2$ as perpetuity
V_{dhs} (Dechow et al. 1999)	Yes	Time-series estimate of persistence of abnormal earnings	12 percent	Not applicable
V_{gm} (Gode and Mohanram 2001)	No	Analysts' earnings forecasts	Fama and French (1997) three-factor, by industry	Perpetual growth rate of 4 percent after $t+2$

TABLE 7
Regression of Future Returns (*SRet36*) on Alternative Measures of Fundamental Value

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	−0.092**	−0.035	0.001	−0.162***	−0.069	−0.082***	−0.082*
	(−2.013)	(−0.894)	(0.001)	(−3.354)	(−1.593)	(−2.808)	(−1.657)
V_f/P	0.112***			0.104**			0.110***
	(2.853)			(2.574)			(2.784)
V_{dhs}/P		0.115***			0.054		
		(2.915)			(1.600)		
V_{gm}/P			0.021**			0.018*	0.001
			(2.030)			(1.711)	(0.056)
B/P				0.122***	0.112***	0.131***	
				(3.990)	(4.384)	(4.086)	
Adj. R^2	0.014	0.009	0.009	0.023	0.013	0.020	0.021
No. of Obs.	18,785	18,785	18,785	18,785	18,785	18,785	18,785
Years	1976–1997	1976–1997	1976–1997	1976–1997	1976–1997	1976–1997	1976–1997

*, **, *** Statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively, (two-tailed).

Regressions are estimated annually and means of the annual estimates are reported. t-statistics are computed as the mean divided by the standard error of the annual coefficient estimates. The Newey and West (1987) procedure is used to adjust for serial correlation induced by overlapping holding periods for *SRet36*.

For computing V_{gm}/P, a firm-year observation is deleted from the sample if (1) one- or two-year-ahead earnings forecast is less than 0, or (2) growth in earnings forecasts over two future years falls in the extreme 0.5 percent of its distribution. In addition, the smallest (largest) 0.5% of the observations of *SRet36*, V_f/P, V_{gm}/P, and V_{dhs}/P are set to the 0.005 (0.995) fractile values of the variables.

SRet36 is the size-adjusted three-year buy-and-hold return beginning July of year *t*, defined as raw buy-and-hold return less the corresponding NYSE/AMEX size-decile index return; V_f/P is analyst forecast-based fundamental value computed using a residual income valuation model, divided by stock price as of the end of June in year *t*; V_{dhs}/P is fundamental value computed using the approach described in Dechow et al. (1999) and Dechow et al. (2001), divided by stock price as of the end of June in year *t*; V_{gm}/P is fundamental value computed using the approach described in Ohlson and Juettner-Nauroth (2000) and Gode and Mohanram (2001), divided by stock price as of the end of June in year *t*; and B/P is the book value in year $t-1$ divided by stock price as of the end of June in year *t*.

significantly in predicting future returns. This result suggests that the ability of V_{dhs}/P to predict returns is primarily due to B/P, and not the *PV(RI)* component, which is the distinguishing feature of the residual income model. Alternatively, if B/P is considered to be a risk factor (Fama and French 1992), then the V_{dhs}/P effect is explained away by an omitted risk factor. Hence, we conclude that V_f/P seems to be a better residual-income-based valuation measure than V_{dhs}/P for predicting future abnormal returns.

Although V_{gm}/P does not embed B/P, for comparison we also report results of a regression of *SRet36* on V_{gm}/P and B/P (see column 6 of Table 7). The incremental contribution of V_{gm}/P is only marginally significant ($t = 1.711$), suggesting that as with V_{dhs}/P, most of the V_{gm}/P effect is explained away by B/P. Finally, we regress *SRet36* on both V_f/P and V_{gm}/P, simultaneously. The results reported in column 7 show that the coefficient on V_f/P is significant ($t = 2.784$) but the coefficient on V_{gm}/P is not significant ($t = 0.056$). These results suggest that V_f/P is better than V_{gm}/P as a predictor of future abnormal returns.[11]

VI. SUMMARY AND CONCLUSIONS

This study provides evidence on the market mispricing and risk explanations for the ability of a residual-income-based value-to-price ratio, V_f/P, to predict future abnormal returns (Frankel and Lee 1998). First, we find that the V_f/P effect is partially concentrated around future earnings announcements, consistent with the mispricing explanation. Second, on using an extensive set of risk proxies suggested in the literature (Gebhardt et al. 2001; Gode and Mohanram 2001), we find that V_f/P is significantly related to some risk proxies. However, after controlling for these risk factors, V_f/P continues to exhibit a significant positive association with future returns. Thus, the V_f/P effect does not seem to be due to V_f/P proxying for risk factors. Overall, we believe that our study provides new evidence that should help readers revise their beliefs about the likelihood of the mispricing and risk explanations for the fundamental value-to-price effect, which has been the subject of some highly influential studies in the accounting and finance literature, namely Frankel and Lee (1998), Lee and Swaminathan (1999), Dechow et al. (1999), and Lee et al. (1999).

REFERENCES

Altman, E. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 22: 589–609.

Ball, R. 1978. Anomalies in relationships between securities' yields and yield surrogates. *Journal of Financial Economics* 5: 103–126.

———, and S. P. Kothari. 1991. Security returns around earnings announcements. *The Accounting Review* 66: 718–751.

———, ———, and J. Shanken. 1995. Problems in measuring portfolio performance: An application to contrarian investment strategies. *Journal of Financial Economics* 38: 79–108.

Beaver, W. H. 2002. Perspectives on recent capital market research. *The Accounting Review* 77: 453–474.

Berk, J. B. 1995. A critique of size related anomalies. *Review of Financial Studies* (Summer): 275–286.

---

[11] We also regress *SRet36* on V_f/P and V_{dhs}/P simultaneously. The coefficients on both the variables are positive and significant ($t = 2.532$ and 2.245, respectively). The significance for the V_{dhs} variable seems to be the manifestation of the column (5) result that suggests V_{dhs}/P predicts returns primarily due to B/P and the column (4) result that suggests that B/P have significant incremental power beyond V_f/P for explaining future returns. Thus, the results from regressing *SRet36* on V_f/P and V_{dhs}/P simultaneously are consistent with our conclusions from the results in Table 7.

Bernard, V., and J. Thomas. 1989. Post earnings announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 22: 445–466.

———, and ———. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13: 305–340.

———, ———, and J. Wahlen. 1997. Accounting-based stock price anomalies: Separating market inefficiencies from risk. *Contemporary Accounting Research* 14: 89–136.

Biddle, G. C., and A. Hunt. 1997. An equity cash flow perspective on the book to market equity ratio's ability to predict stock returns. Working paper, Hong Kong University of Science and Technology.

Brennan, M., N. Jegadeesh, and B. Swaminathan. 1993. Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies* 6: 799–824.

Chari, V., R. Jagannathan, and A. Ofer. 1988. Seasonality in security returns: The case of earnings announcements. *Journal of Financial Economics* 21: 101–122.

Dechow, P., A. Hutton, and R. Sloan. 1999. An empirical assessment of the residual income valuation model. *Journal of Accounting and Economics* 26: 1–34.

———, ———, L. Meulbroek, and R. Sloan. 2001. Short-sellers, fundamental analysis and stock returns. *Journal of Financial Economics* 61: 77–106.

Fama, E., and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 46: 427–466.

———, and ———. 1997. Industry costs of equity. *Journal of Financial Economics* 43: 153–193.

Frankel, R., and C. M. C. Lee. 1998. Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics* 25: 283–320.

Freeman, R. 1987. The association between accounting earnings and security returns for large and small firms. *Journal of Accounting and Economics* 9: 195–228.

Gebhardt, W., C. M. C. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39: 135–176.

Gode, D., and P. Mohanram. 2001. Implied cost of capital. Working paper, New York University, New York.

Herzberg, M. M. 1998. Implementing EBO/EVA analysis in stock selection. *The Journal of Investing*: 45–53.

Keim, D. B. 1988. Stock market regularities: A synthesis of the evidence and explanations. In *Stock Market Anomalies*, edited by E. Dimson. Cambridge, U.K.: Cambridge University Press.

Kothari, S. P. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31: 105–232.

La Porta, R., J. Lakonishok, A. Shleifer, and R. Vishny. 1997. Good news for value stocks: Further evidence on market efficiency. *Journal of Finance* 52: 859–874.

Lee, C. M. C. 1996. Measuring wealth. *CA Magazine* (April): 32–37.

———, J. Myers, and B. Swaminathan. 1999. What is the intrinsic value of the Dow? *The Journal of Finance* 54: 1693–1741.

———, and B. Swaminathan. 1999. Valuing the Dow: A bottom-up approach. *Financial Analyst Journal* (September/October): 4–23.

Lo, K., and T. Lys. 2000. The Ohlson model: Contribution to valuation theory, limitations and empirical applications. *Journal of Accounting, Auditing and Finance* 15: 337–367.

Merton, R. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42: 483–511.

Newey, W., and K. West. 1987. Hypothesis testing with efficient method of moments estimation. *International Economic Review* 28: 777–787.

Noreen, E. 1989. *Computer Intensive Methods for Testing Hypotheses*. New York, NY: John Wiley & Sons.

Ohlson, J. 1990. A synthesis of security valuation theory and the role of dividends, cash flows and earnings. *Contemporary Accounting Research* 6: 648–676.

———. 1995. Earnings, book values and dividends in security valuation. *Contemporary Accounting Research* 11: 661–687.

———, and B. Juettner-Nauroth. 2000. Expected EPS and EPS growth as determinants of value. Working paper, New York University, New York.

Skinner, D., and R. Sloan. 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7: 289–312.

THE ACCOUNTING REVIEW
Vol. 78, No. 2
2003
pp. 397–428

Heterogeneous Investment Opportunities in Multiple-Segment Firms and the Incremental Value Relevance of Segment Accounting Data

Peter F. Chen
Guochang Zhang
Hong Kong University of Science and Technology

ABSTRACT: Applying a real-options-based valuation approach, we develop and test a model that addresses the incremental value relevance of segment data beyond firm-level accounting data. Prior studies (e.g., Zhang 2000; Biddle et al. 2001) show that equity valuation requires accounting data (in part) because accounting provides signals that guide capital investments underlying value creation. In this study, we establish that the usefulness of segment data beyond aggregate data relates to heterogeneity of investment opportunities across segments, caused by *divergences* of segment profitability and growth potential. Empirical results are consistent with the model's predictions. We also assess the magnitude of the valuation impact of segment information relative to that of firm-level information.

Keywords: *segment accounting data; incremental value relevance; real options; divergence of profitability; segment growth opportunity.*

Data Availability: *Data used in this study are available from public sources.*

I. INTRODUCTION

Valuation of a firm with multiple business components requires information not only about the firm's overall operations, but also about the operations of individual components (segments) within the firm. Analysts and investors view segment data as vital and essential for investment analysis,[1] and they demand segment data that enhance

---

[1] The Association for Investment Management and Research (AIMR) in its position paper (AIMR 1993, 59-60) state, "[Segment information] is vital, essential, indispensable, and integral to the investment analysis process.... Different segments will generate dissimilar streams of cash flows to which are attached disparate risks and which bring about unique values. Thus, without disaggregation, there is no sensible way to predict the overall amounts, timing, or risks of a complete enterprise's future cash flows. There is little dispute over the analytic usefulness of disaggregated financial data."

We appreciate the comments of Daniel Thornton (Associate Editor), two anonymous referees, and workshop participants at the Chinese University of Hong Kong, the Hong Kong University of Science and Technology, and the 2001 HKUST Summer Symposium. We also thank Li Xia for research assistance. We gratefully acknowledge the financial support from Research Grants Council of Hong Kong (HKUST6023/00H). All errors are our own.

Editor's note: This paper was accepted by Daniel B. Thornton, *Ad Hoc* Editor.

Submitted February 2001
Accepted November 2002

their ability to predict management actions that affect future cash flows (SFAS No. 131, FASB 1997). In response to calls from the investment community for more disaggregated information, standard setters and regulators worldwide have recently made considerable efforts to expand segment disclosure requirements (e.g., SFAS No. 131; No. 14R, International Accounting Standards Committee [IASC] 1997) and to enforce the rules.[2] Despite the importance of segment data to users and regulators, there has been limited academic research to formally examine their role for valuation. In this study, we employ a real-options-based model to address the following question: Given aggregate data that firms are already required to report, how is equity value related to the information conveyed *incrementally* by segment data?[3] We first develop a model to show how equity value depends on both firm-level and segment accounting information, and then empirically tests the model's predictions.

It is well known that firm value reflects the value of real options, i.e., opportunities to expand or contract the scale of operations (e.g., Myers 1977; Kester 1984; de Matos 2001). Real options exist in virtually all types of business operations, and they often constitute a significant portion of firm value. Accounting research has previously examined the implications of firm-level real options or, equivalently, treated firms as single-segment operations (e.g., Hayn 1995; Berger et al. 1996; Burgstahler and Dichev 1997; Zhang 2000; Yee 2000; Biddle et al. 2001), but the underlying idea can be readily extended to firms with multiple business segments. In this study, we recognize that in a multiple-segment firm, different segments generally face different investment opportunities (due to different external market conditions), and have different abilities to exploit external opportunities (due to differential abilities to manage business). In a model that incorporates segment-level real options, we decompose equity value into (1) the part explained by aggregate firm-level accounting data, and (2) an incremental component attributable to differences across segments in operating profitability and investment opportunities as conveyed by segment-level data. This characterization enables us to evaluate the importance of segment-level, vs. firm-level, data in value determination and to explain cross-sectional variations in the relative importance of segment data as driven by underlying economic conditions.

To examine the incremental value relevance of segment data, we first consider the case where a firm's segments have equal profitabilities and growth opportunities. Under these special conditions, segments have identical real options, and firm value can be determined solely by firm-level data. When segments have divergent profitabilities, valuation requires not only firm-level accounting data to convey about firm overall operations, but also segment-level data to convey differences among segments. A key valuation construct is divergence of profitability (DOP), which indicates the differential abilities of segments in exploiting investment opportunities. The incremental value relevance of segment data increases with DOP, because DOP is positively associated with the heterogeneity of investment decisions across segments.

---

[2] SEC enforcement cases demonstrate that regulators take a serious stand in enforcing segment disclosure rules. On August 15, 1988, for example, the SEC imposed a $1 million penalty on Sony Corporation for failing to report Sony Pictures as a separate industry segment, but instead reporting the combined results of Sony Picture and Sony's profitable music business as a single "entertainment" segment (SEC 1998). This reporting policy obscured the losses sustained by Sony Pictures. (We thank the associate editor for bringing this to our attention.)

[3] We emphasize the *incremental* role of segment data in this study because in practice segment data are reported as additional disclosures to supplement firm-level financial data. This is consistent with the incremental usefulness of segment data as emphasized by both users and standard setters (see, for example, AIMR 1993; AICPA 1994).

More generally, when a firm's segments operate in industries or markets that have different growth opportunities, equity value depends not only on the divergence of profitability, but also on the distribution of growth opportunities across segments. When most of a firm's growth opportunities lie in segments that are relatively profitable, a situation advantageous to the firm in exploiting external opportunities, DOP increases total firm value (relative to where segments have equal profitability). On the other hand, when most of the firm's growth opportunities lie in segments that are relatively unprofitable, a situation disadvantageous to the firm in exploiting opportunities, DOP reduces firm value.

We perform empirical analysis on data from the Compustat industry segment file for the period 1986–1997, and obtain the following results. First, controlling for firm-level accounting information, DOP exhibits a significantly positive effect on equity value. Second, the incremental value effect of a given DOP varies with firm overall profitability and growth opportunity as predicted. Third, the distribution of segment growth opportunities within a firm is important in determining the value effect of DOP; the incremental value effect of DOP is positive if growth opportunities are more highly concentrated in segments that are relatively profitable, and is negative if the opposite is the case. Overall, the empirical results are consistent with qualitative predictions of the model.

The magnitude of the value effect of DOP is shown to be *economically* significant. For example, on a broad-based sample, the estimated equity value change associated with a change of 0.10 in DOP (approximately 2 standard deviations in partitioned samples) ranges from +6.7 percent to –6 percent, depending on how growth opportunities are distributed within the firm. On the other hand, the incremental explanatory power of DOP is generally small, possibly due to measurement errors in reported segment data (see Givoly et al. 1999).

Our study suggests that when valuing multiple-segment firms, firm-level data need to be supplemented by information that reflects differences in segment profitability and growth opportunities, the underlying causes for segments to pursue heterogeneous investment decisions. This is in line with the arguments for disclosure of disaggregated information made by analysts and standard setters. In addition, by establishing a link between the valuation relevance of segment data and their use in managerial decisions, our study suggests that it is beneficial to organize external reporting of segment data in a way consistent with the internal use of such data by managers. In this sense, our analysis supports both the management approach underlying SFAS No. 131 and the SEC requirement to organize the Management Discussion and Analysis (MD&A) in a way that portrays the firm "through the eyes of management." Also, our result that segments can be aggregated when they have similar profitabilities and growth opportunities is consistent with the aggregation criteria recommended in SFAS No. 131.[4]

The remainder of the paper is organized as follows. Section II reviews prior research. Section III develops a real-options-based valuation model for multiple-segment firms, which leads to testable hypotheses. Section IV describes variable measurement and data. Section V reports empirical results, and Section VI discusses implications and concludes the paper.

---

[4] Specifically, SFAS No. 131 states that separate reporting of segment information will not add significantly to an investor's understanding of an enterprise if its operating segments have characteristics so similar that they can be expected to have essentially the same future prospects (FASB 1997, 37).

II. PRIOR LITERATURE

A number of prior studies have examined the role of segment accounting data in valuation, with different approaches employed.[5] Foster (1975) compares market performances of portfolios formed by aggregate-earnings news vs. portfolios formed by sub-earnings news (earnings news is classified into negative or positive changes), and finds systematic differences in portfolio returns. This evidence, however, may merely indicate differential stock market reactions to the different *magnitudes* of firm-level earnings (as the *magnitude* of total earnings is expected to differ systematically across portfolios), and may not necessarily suggest that earnings disaggregation, per se, provides relevant information.

More closely related to our study is Wysocki (1998) who adopts a real-options framework to predict how the valuation consequence of segment earnings differs for different segments. Based on a sample of two-segment firms, Wysocki finds that (1) earnings coefficients are higher for profit (vs. loss) segments, for more profitable (vs. less profitable) segments, and for the primary-industry segment whose growth opportunity exceeds that of the secondary-industry segment; and (2) the incremental explanatory power of earnings is higher for profit segments than for loss segments. These findings indicate that the usefulness of segment-level data depends on whether segments differ in their profitabilities and growth opportunities.

Basu et al. (1999) model and test the usefulness of segment earnings in relation to segment earnings attributes (namely, differences in persistence, variances, and correlation). They run separate regressions of stock returns on firm-level data and on segment data. Based on comparison of regression coefficients on segment earnings with that of aggregate earnings, they find that the usefulness of segment data depends on the above earnings attributes. Givoly et al. (1999) employ similar regressions but their comparison focuses on regression R^2s. They find that increases in R^2s are small when component segment earnings, instead of firm-level earnings, are used to explain returns, a result also reported by Basu et al. (1999).

The above studies treat segment earnings as alternative, rather than incremental, explanatory variables to firm-level earnings. They do not address the *incremental* role of segment data in a valuation model that already incorporates firm-level data, or how segment data help to "fine tune" equity valuation. This issue is pertinent to practice because segment data are provided as supplementary disclosures, and regulators and users evaluate the usefulness of segment data given firm-level data.

One study that incorporates segment data to supplement aggregate earnings is Tse (1989), who finds that the percentage of total earnings derived from a "high-" (vs. "low-") growth industry is additionally value-relevant. Tse (1989), however, does not develop a theoretical model to show endogenously what types of segment data matter incrementally and what their predicted effects are; his analysis focuses on one specific type of segment data.

Our study contributes to further understanding the usefulness of segment data, both theoretically and empirically. Our model shows why segment data incrementally matters, how value-relevant information is extracted, and how its incremental valuation impact varies with divergence of segment profitability, firm overall profitability, overall growth opportunity, and the distribution of growth opportunities within the firm. The study also provides

---

[5] In addition, there are other studies examining issues that are indirectly related to valuation such as the effect of industry/geographical segment data on investors' expected returns and risk assessments (see Pactor [1993] for a survey) or on earnings forecast accuracy (Lobo et al. 1998).

an assessment of the magnitude of the valuation impact of segment information relative to that of firm-level information.

III. THE VALUATION MODEL AND PREDICTIONS

This section develops a valuation model for multiple-segment firms to examine the incremental valuation role of segment data. The model is an extension of Zhang (2000), which introduces real options into the valuation framework of Ohlson (1995) and Feltham and Ohlson (1995, 1996). The notion that equity value reflects real options is empirically supported by Berger et al. (1996), Hayn (1995), Burgstahler and Dichev (1997), Wysocki (1999), and Collins et al. (1999).

A Single-Segment Firm Valuation Model

According to Zhang (2000), and consistent with Burgstahler and Dichev (1997), equity value can be viewed as equal to the value from existing operations plus the value of options to grow or to contract the operating scale, and it is a function of the firm's profitability and growth opportunities. For expositional clarity, we present a simplified version of Zhang's (2000) model, used as the basis for developing our analysis for multiple-segment firms.[6]

Let V be the intrinsic value of an all-equity financed firm, A the amount of asset stock currently employed in operations, E the current-period earnings, g the growth opportunity, i.e., the percentage by which the scale of operations (as measured by assets) can grow, and c the earnings capitalization factor.[7] Define $q \equiv E/A$ as the firm's current operating profitability. Following Zhang (2000), the value of a firm with a growth and an adaptation option can be expressed as a function of q as follows:

$$V = AP(q) + cE + AgC(q) = A[P(q) + cq + gC(q)], \tag{1}$$

where $P(q)$ is the (put) option that allows the firm to adapt assets, and $C(q)$ the (call) option that allows the firm to grow, both of which are normalized by assets. The options' values depend on current profitability (q) as long as current profitability is informative about future profitability and hence affects the likelihood that the options will be exercised. Embedded in these options are two critical profitability levels (the "exercise prices"), denoted q_a^* and q_g^*, $q_a^* < q_g^*$, such that the firm will adapt its assets if future profitability is below q_a^*, expands operations if future profitability is above q_g^*, and maintains the existing scale in between. A property important for the analysis below is that equity value V as given by Equation (1) is an increasing and convex function of q, with the degree of convexity increasing with growth opportunity g.

Valuation of Multiple-Segment Firms

Consider a firm with two segments, labeled segments 1 and 2, where each segment can choose to adapt its resources or to grow, irrespective of the other segment's decision. Following the above analysis, each segment has a value expression analogous to Equation (1). Combining the two segments, we obtain firm value as:[8]

---

[6] See Zhang (2000) for the original version of the model and the details of derivation.

[7] For notational simplicity and where no confusion arises, we omit time subscripts on variables.

[8] For simplicity, we assume away any interactive effect (such as synergy) between segments. Our results will hold quantitatively even in the presence of such effects. What drives our results is a feature generic to all multiple-segment firms (with or without synergies), that there exists at least some degree of independence in operating decisions among segments; it is this independence that gives rise to segment-level real options.

$$V = [A_1P(q_1) + cE_1 + A_1g_1C(q_1)] + [A_2P(q_2) + cE_2 + A_2g_2C(q_2)]$$
$$= [A_1P(q_1) + A_2P(q_2)] + c[A_1q_1 + A_2q_2] + [A_1g_1C(q_1) + A_2g_2C(q_2)], \quad (2)$$

where, by definition, $q_1 = E_1/A_1$ and $q_2 = E_2/A_2$, and subscript 1 or 2 identifies a segment.

Let A and E be the firm's total assets and current-period earnings, respectively, with $A = A_1 + A_2$ and $E = E_1 + E_2$, and let $q = E/A$ be the firm's overall profitability. Then, q is a weighted average of segment profitabilities, $q = w_1q_1 + w_2q_2$, with weights $w_1 = A_1/A$ and $w_2 = A_2/A$.

Due to the option terms, valuation function (2) is a nonlinear function of segment assets (A_1, A_2), segment profitability (q_1, q_2), and growth opportunities (g_1, g_2). This nonlinearity generally makes it inappropriate to replace segment earnings and book values in Equation (2) with firm-level data, suggesting that firm-level accounting data alone do not provide sufficient information for valuation.

Of special interest are situations where segment data can be aggregated in Equation (2). This occurs when the segments have equal profitability ($q_1 = q_2$) and equal growth opportunities ($g_1 = g_2$). Under these conditions, segments have identical real options, and Equation (2) simplifies to:

$$V = AP(q) + cE + AgC(q), \quad (3)$$

equivalent to the value expression for a single-segment firm, Equation (1), where firm overall growth opportunity g coincides with g_1 and g_2 in this case.

In general, when segments differ in profitability and growth opportunities, segment-level data are necessary to supplement firm-level data to convey additional relevant information. Below, we examine the incremental relevance of segment data first for the case where segments have equal growth opportunities and then extend the analysis to more general situations.

The Incremental Relevance of Segment Data When Segments Have Equal Growth Opportunities ($g_1 = g_2 = g$)

The segments have same growth opportunities if, for example, they operate in the same (or similar) market environments. Designate the less profitable segment as segment 1 and the other as segment 2, i.e., $q_1 \leq q_2$. Define $DOP \equiv q_2 - q_1 \geq 0$ as the divergence of profitability between the segments. Let $DOP = 0$ (i.e., $q_2 = q_1$) be a benchmark case, with firm value given by Equation (3).[9] Figure 1 illustrates how *DOP* affects firm value *incremental to* firm-level data under a convex valuation function.[10]

Curve AB in Figure 1 represents value, normalized by assets, as a function of profitability; it applies to each individual segment and also to the whole firm in the benchmark case ($DOP = 0$). When both segments' profitabilities (and hence firm overall profitability) equal q, the normalized firm value and that of each segment are all indicated by point M (on curve AB). As segment profitabilities move apart to q_1 (for segment 1) and q_2 (for segment 2), while holding firm overall profitability unchanged at q, the (normalized) value of segment 1 moves to point J and that of segment 2 to K. Based on Equation (2), it can

---

[9] In the empirical analysis below, the definition of *DOP* is generalized to firms with more than two segments.

[10] The valuation properties are conveniently illustrated with graphs. Alternatively, the properties can be shown analytically, but the exposition would be much more complicated. The analytical derivations are available upon request.

FIGURE 1

Effect of Divergence of Profitability on Firm Value When $g = g_2$

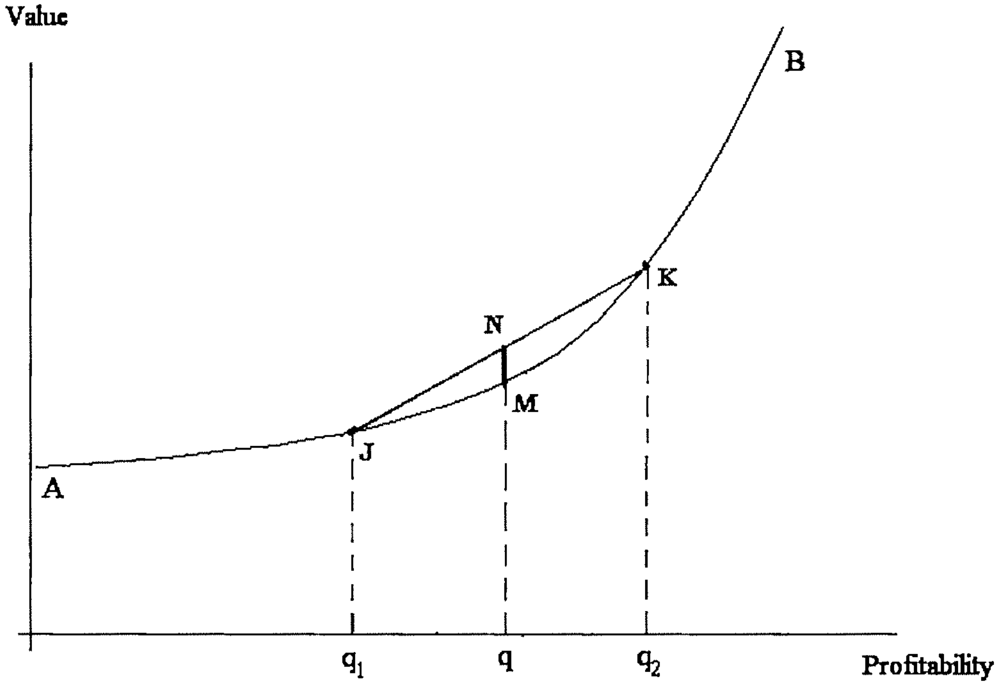


Figure 1 illustrates how divergence of segmental profitability ($DOP = q_2 - q_1$) affects firm value when segments 1 and 2 have the same growth opportunities ($g_1 = g_2$). The vertical axis is *value* (normalized by assets), and the horizontal axis is *profitability*. Curve AB represents normalized value as a function of profitability for each segment and also for the firm as a whole in the benchmark case $DOP = 0$. Point M indicates the firm value if the profitability of both segments (and hence the firm) equals q. As the two segments diverge in profitability, $DOP = q_2 - q_1 > 0$, while holding firm overall profitability unchanged at q, the value of the less profitable segment (labeled segment 1) decreases, as given by point J, and that of segment 2 increases, given by point K. The normalized firm value, a weight average of segment values, is given by point N. N lies above M, due to convexity of AB. The vertical distance between N and M represents the incremental value attributed to DOP.

be shown that the normalized firm value is a weighted average of normalized segment values,[11] which is now indicated by point N. Given that the valuation function is convex, N lies above M, suggesting that, relative to the benchmark case of $DOP = 0$, firm value is greater when segment profitabilities move apart.[12] Distance NM represents the *incremental* value effect of DOP, denoted Δv. It is obvious from Figure 1 that given overall profitability q, Δv increases with DOP. Our first hypothesis, stated in alternative form, is:

---

[11] From Equation (2), $V/A = w_1(V_1/A_1 + w_2(V_2/A_2)$ where $V_1/A_1 = P(q_1) + cq_1 + g_1C(q_1)$ and $V_2/A_2 = P(q_2) + cq_2 + g_2C(q_2)$ are normalized segment values, and weights $w_1 = A_1/A$ and $w_2 = A_2/A$.

[12] Firm profitability q is a weighted average of q_1 and q_2, and the normalized firm value is a weighted average of normalized segment values, where the same set of weights applies to both averages. A positive value effect of DOP follows from Jensen's Inequality that for a strictly convex function $f(x)$ and constant w_1, $0 < w_1 < 1$, we have $f(w_1x_1 + (1 - w_1)\,x_2) > w_1\,f(x_1) + (1 - w_1)f(x_2)$.

H1: Divergence of segment profitability (*DOP*) has a positive incremental effect on equity value, given firm overall profitability q.[13]

However, the magnitude of Δv associated with a *given DOP* depends on how convex the valuation function is around profitability q. The more convex the valuation function around q, the greater the incremental value impact of a given *DOP*. In one extreme where value is linear in profitability, the effect of *DOP* vanishes.

According to the established option theory (e.g., Hull 2000), option value (either put or call) is generally convex, but the degree of convexity is the highest where the option is "at the money" (i.e., where the price of the underlying asset equals the exercise price). Moving away in either direction from this point, option value becomes less convex and gradually approaches linearity.[14] In our valuation context, these properties suggest that the valuation function is most convex near q_a^* and q_g^*, the critical profitability levels for adaptation and growth decisions. The two options jointly induce the following behavior of Δv: For a given *DOP*, Δv reaches a maximum when firm profitability q is near q_a^* and also near q_g^* (where the valuation function is most convex); Δv increases with q to the left of q_a^* and decreases with q to the right of q_g^*, approaching zero in both extremes where $q << q_a^*$ or $q >> q_g^*$. Between q_a^* and q_g^*, Δv decreases with q in the region near q_a^* (where the effect of the adaptation option dominates that of the growth option) and increases with q in the region near q_g^* (where the effect of growth option dominates), forming a valley between the two maximums.

Figure 2 shows the simulation results on how Δv behaves in relation to q for a range of different *DOP* values. It confirms the qualitative trend predicted in the above discussion. As shown in the figure, the parameter values underlying the simulation are comparable to those of average firms in our empirical sample. The figure gives a general indication that the magnitude of the incremental effect of *DOP*, while not comparable to that of firm-level information, would be (on average) economically important for firms and investors. Furthermore, this effect varies considerably across firms. (This general point also applies to the additional simulation in Figure 4 below.)

H2: In the extreme low-q region, Δv is close to zero. As q increases, Δv first experiences an increasing trend, followed by a decreasing trend, and then another increasing trend followed by a decreasing trend, and it eventually approaches zero in the extreme high-q region.

In Figure 2, the peaks on the right are induced by the growth options. These curves are based on a particular value of g (with different values of *DOP*). From Equations (2) and (3), growth opportunities have a multiplying effect on value; therefore, as g increases, the incremental value effect Δv of a *given DOP* also increases. In the special case where $g = 0$, there are no growth opportunities, and the peaks on the right would disappear in Figure 2.

H3: The incremental value (Δv) of a given *DOP* increases with growth opportunity g.

---

[13] This hypothesis is later shown to hold under more general conditions of nonnegative correlation between segment profitability and growth opportunity.

[14] This is because as the option moves further into the region of "in the money" or the region of "out of the money," the likelihood that the option will be exercised becomes less sensitive to changes in the underlying asset price (the option-holder is more sure about what his eventual decision will be), and this causes option value to be less convex (i.e., more close to linearity).

FIGURE 2

Simulation Result on the Effect of DOP on Firm Value When $g_1 = g_2$

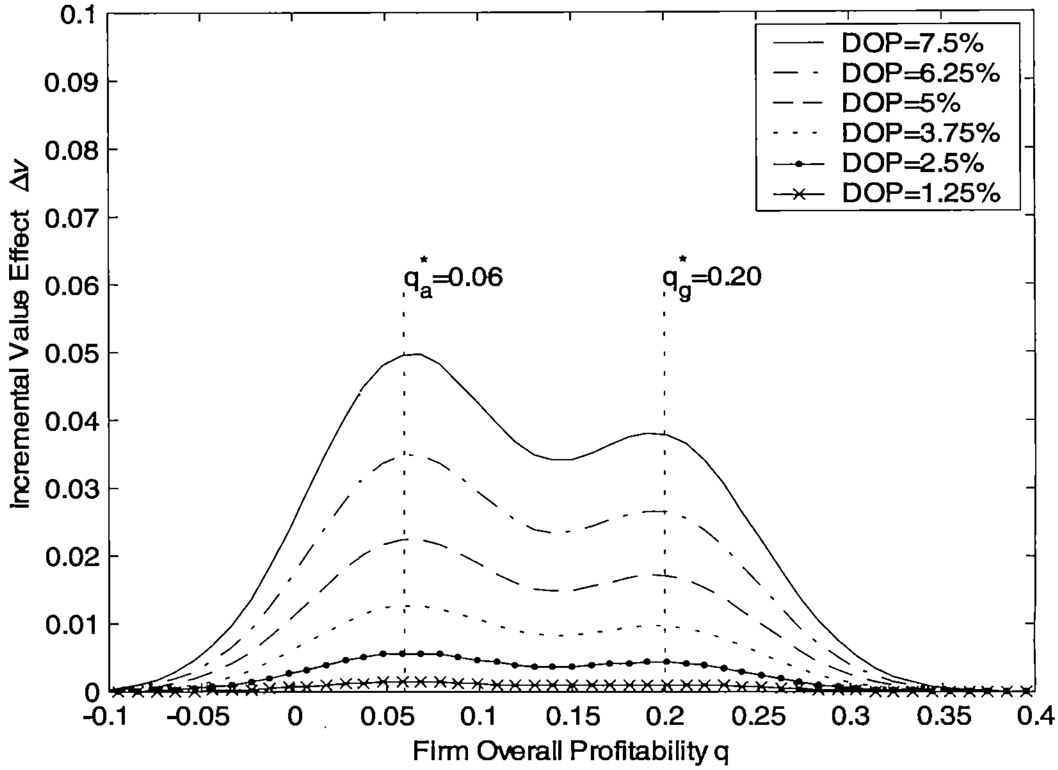


Figure 2 presents the simulation result on the incremental value effect of divergence of segmental profitability (*DOP*) when segments 1 and 2 have the same growth opportunities ($g_1 = g_2$). The vertical axis (Δv) represents the increase or decrease in firm value (normalized by assets), relative to the benchmark firm value given by Equation (3) (where $DOP = q_2 - q_1 = 0$), that is attributed to *DOP*, and the horizontal axis (q) is firm overall profitability. Each curve shows the relation between Δv and q for a particular value of *DOP*. Growth opportunities are set to be $g_1 = g_2 = 0.75$ in the figure. The results are qualitatively the same for other parameter values.

For comparison, the mean analyst forecast of earnings growth rate is 12 percent annually for our overall empirical sample (which implies a cumulative growth of 76 percent if this lasts for five years). The mean *DOP* is 7 percent for the overall sample, and it ranges from 3 percent to 11 percent for the subsamples employed in our empirical analysis.

The Incremental Relevance of Segment Data When Segments Have Different Growth Opportunities ($g_1 \neq g_2$)

When segments have different growth opportunities, segment data relate to equity value in a more complex way because equity value also depends on the divergence of growth opportunity between the segments. To isolate the effect of growth opportunities, we distinguish firms with high profitability from those with low profitability. Our analysis below focuses on the high-profitability region where growth options are an important part of firm value. In the region where profitability is low (relative to q_g^*), growth options have little relevance to valuation (deep "out of the money"), and Δv behaves similarly to that for the case of $g_1 = g_2$ above (Figure 2). Namely, Δv is negligible when $q << q_a^*$, and, as q increases, Δv first increases to reach a maximum near $q = q_a^*$ and then begins to decline.

In the high-profitability region, the incremental value effect of *DOP* is illustrated by Figure 3, which generalizes the special case of $g_1 = g_2$ in Figure 1. With $g_1 \neq g_2$, two value curves are needed to depict the relation between segment value and profitability: The more convex curve, AB, corresponds to the segment with higher growth opportunity, whereas the less convex AC corresponds to the segment with lower growth opportunity (all values normalized by assets); the segments' values are approximately the same when profitability is low. As before, our benchmark is the special case of $q_1 = q_2 = q$, with firm value given by:

$$V = (A_1 + A_2)P(q) + c(A_1 + A_2)q + A_1g_1C(q) + A_2g_2C(q). \tag{3'}$$

The normalized firm value, a weighted average of normalized segment values (see footnote 11), in the benchmark is indicated by point M, which lies between two segments' values. As segment profitabilities diverge, while keeping firm profitability unchanged at q, segment values move along their respective value curves. There are two distinct scenarios.

Scenario (i): The more profitable segment has higher growth opportunity ($g_1 < g_2$).

When $g_1 < g_2$, AB (the more convex curve) corresponds to segment 2 and AC to segment 1. Starting from the benchmark where both segments have profitability q, let segment profitabilities diverge to q_1 (for segment 1) and q_2 (for segment 2) but without changing overall profitability q. The value of segment 1 moves to J, and that of segment 2 to K. The normalized firm value is given by N, located above the benchmark value at M. In this case, as in the case of $g_1 = g_2$ above, *DOP* has a positive incremental effect on firm value, with the magnitude of the effect measured by distance NM.

Scenario (ii): The less profitable segment has higher growth opportunities ($g_1 > g_2$).

When $g_1 > g_2$, Figure 3 is reinterpreted so that now curve AB corresponds to segment 1 and AC to segment 2. The value of segment 1 is given by point J′, and that of segment 2 by K′. The firm value is given by N′, located below M. In this case, and contrary to the case of $g_1 \leq g_2$, *DOP* has a negative incremental value effect, with the magnitude of the effect represented by distance MN′.[15]

Intuitively, the incremental value of *DOP* depends on how growth opportunities are distributed within the firm. A more profitable segment is more capable of generating value from given investment opportunity than a less profitable segment. Thus, the situation is advantageous to the firm if most of its opportunities lie with the relatively profitable segment (as in Scenario (i) above), but is disadvantageous if the opposite is true (as in Scenario (ii) above). This explains why the value effect of *DOP* is positive in the former scenario and negative in the latter, relative to benchmark value given by Equation (3′).

The generality of the above predictions are confirmed by the simulation result in Figure 4. The *difference* in growth opportunity between segments is defined as $\Delta g = g_2 - g_1$.[16] Each curve shows how Δv changes with q for a given Δg. Two observations can be made. First, Δg matters to valuation in the high-profitability region, but not in the low-profitability region where the behavior of Δv is similar to the case of $g_1 = g_2$ shown in Figure 2. Second, in the high-profitability region, Δv increases with Δg, and the sign of Δv is closely related to that of Δg, as is illustrated in Figure 3.

---

[15] The result illustrated in Figure 3 is a general theoretical result. It is confirmed by the simulation result in Figure 4 below.

[16] In the empirical analysis below, the definition of Δg is extended to general multiple-segment firms, which takes into account differences in segment size.

FIGURE 3
Effect of Divergence of Profitability on Firm Value When $g_1 \neq g_2$

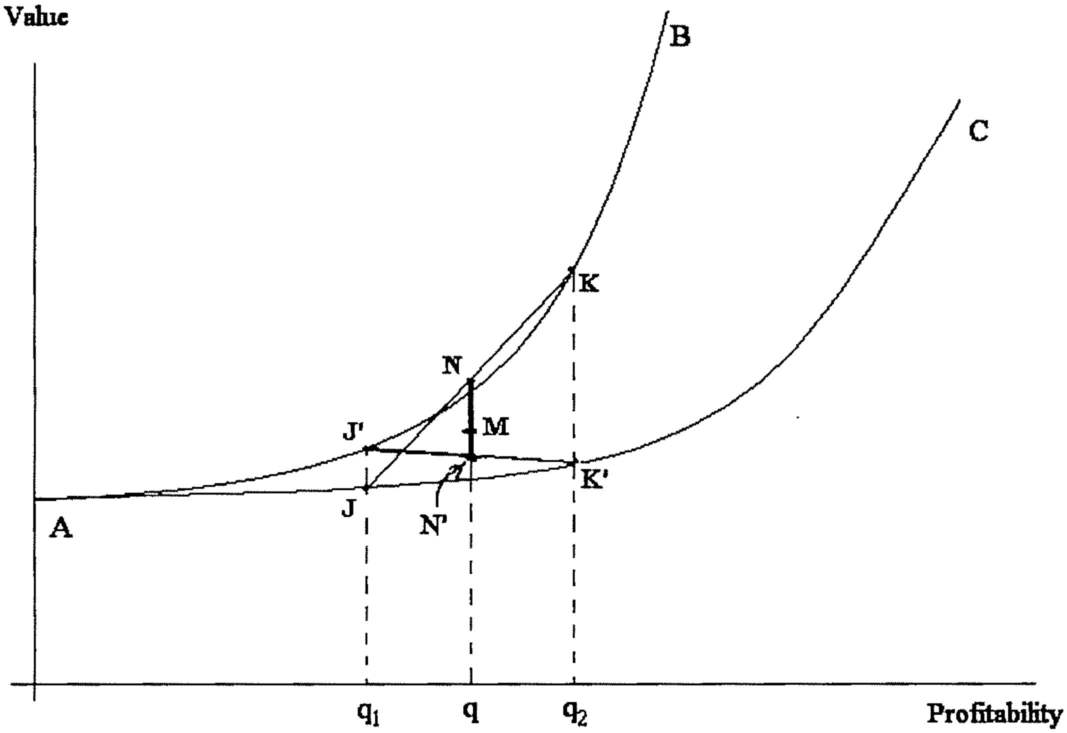


Figure 3 illustrates the incremental value effect of divergence of profitability (*DOP*) and how this effect is influenced by the difference in growth opportunities between segments (Δg). The vertical axis is *value* (normalized by assets) and the horizontal axis is *profitability*. Curve AB corresponds to the segment with higher growth opportunity, whereas AC corresponds to the segment with lower growth opportunity. When the profitability of both segments (and hence of the firm) equals q, the normalized firm value (a weighted average of segment values) is given by point M, which lies between the two value curves. When segment profitabilities change, while holding firm overall profitability unchanged, the firm value changes, as illustrated in the following two scenarios.

***Scenario (I)**: The more profitable segment has higher growth opportunity.* The normalized value of the less profitable segment (segment 1), at profitability q_1, is given by point J, and that of the more profitable segment at profitability q_2 is given by K. The (normalized) firm value is given by N, which lies above M. Distance NM represents the incremental value effect of *DOP* (Δv), relative to benchmark value given by Equation (3'), and it is positive.

***Scenario (II)**: The more profitable segment has less growth opportunity.* The value of the less profitable segment (segment 1) is given by point J', that of the more profitable segment is given by point K', and the normalized firm value is given by N', which lies below M. Distance MN' represents the incremental value effect of *DOP* (Δv), relative to Equation (3'), and it is negative.

H4: In the high-profitability region, the incremental value of *DOP* (relative to the benchmark of $DOP = 0$), Δv, increases monotonically with Δg; $\Delta v > 0$ if $\Delta g > 0$, and $\Delta v < 0$ if $\Delta g < 0$.

IV. EMPIRICAL MEASURES AND DATA

We obtain segment data from the Compustat industry-segment file for period 1986–1997, firm-level data from the Compustat annual data file, and stock prices from the CRSP

FIGURE 4
Simulation Result on the Effect of DOP on Firm Value When $g_1 \neq g_2$

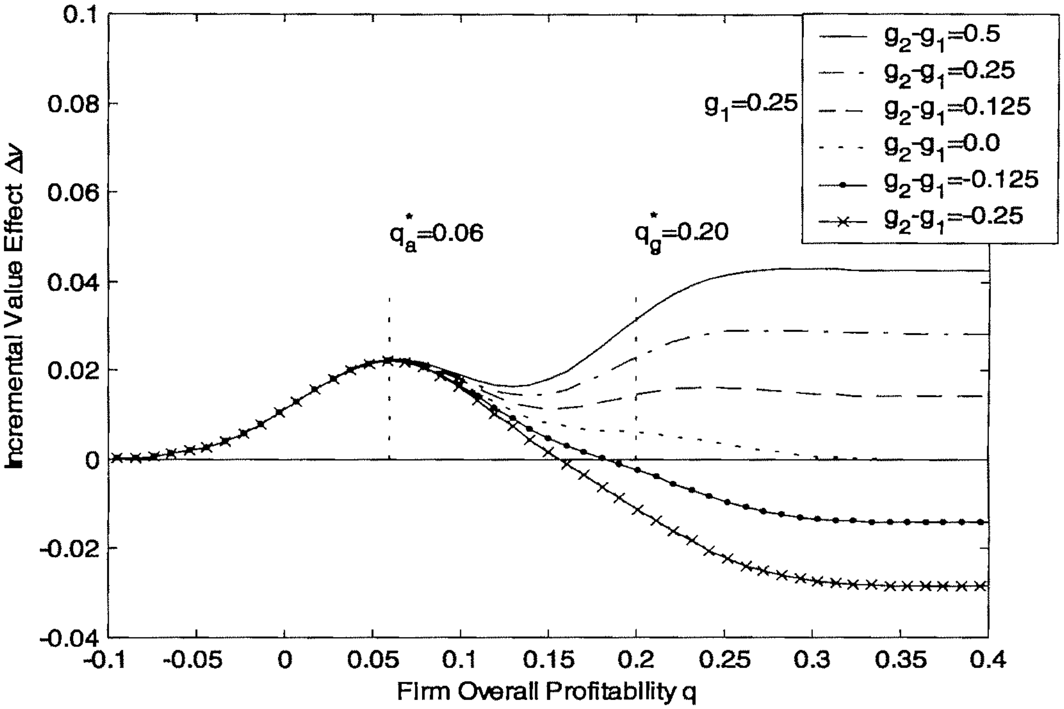


Figure 4 presents the simulation result on how the difference in growth opportunities between segments Δg (which equals $g_2 - g_1$, assuming that the segments have equal amount of assets) influences the incremental value effect *DOP*. The vertical axis (Δv) represents the increase or decrease in firm value, relative to the benchmark value given by Equation (3′), that is attributed to *DOP*, and the horizontal axis (q) represents firm profitability. Each curve corresponds to a particular value of $\Delta g = g_2 - g_1$. $DOP = 5$ percent and $g_1 = 0.25$ for all curves. The results are qualitatively the same for other parameter values.

For comparison, the mean of Δg ranges from -7 percent to 7 percent in terms of *annual* earnings growth rate for various subsamples used in our empirical analysis. If the annual growth rate is 4 percent for segment 1 (implying a total growth of 22 percent if this lasts for 5 years), then the mean cumulative growth difference over a five-year period between two segments, as implied by above mean annual differences, would range from -36 percent to 47 percent. The mean *DOP* is 7 percent for the overall sample, and it ranges from 3 percent to 11 percent for the subsamples employed in our empirical analysis.

database. Equity value at year t (V_t) is measured by market value of common shares, equal to the number of shares outstanding (item #25) times the stock price three months after fiscal year-end. Beginning equity book value B_{t-1} (item #60) is used as a deflator for V_t. Firm overall profitability (q_t) is measured by return on book equity, i.e., earnings (E_t) divided by beginning book equity (B_{t-1}), where E_t is the firm's aggregate earnings before extraordinary items available for common shares (item #237), and B_{t-1} proxies for the capital contributed by equityholders plus retained earnings (which corresponds to variable A_{t-1} in the above model with all-equity financing).

At the segment level, we use return on operating assets, defined as segment operating earnings OE_t (industry-segment file item #2) divided by beginning book value of segment identifiable assets OA_{t-1} (industry-segment file item #5), to proxy for segment profitability; that is, for segment i in period t, $q_{it} = OE_{it}/OA_{i,t-1}$.[17]

Extending the notion of divergence of profitability (*DOP*) defined in Section III, taking into account the effect of different segment sizes, we now introduce a more general definition of *DOP* applicable to empirical data as:

$$DOP = \sum_{i=1}^{k} |q_{it} - \bar{q}_t| * w_{it} \tag{4}$$

where k is the number of segments of the firm, $q_{it} = OE_{it}/OA_{i,t-1}$ is profitability of segment i, $\bar{q}_t = \Sigma_{i=1}^{k} w_{it}q_{it}$ is the weighted average of segment profitability, with weights (w_{it}) defined in proportion to identifiable segment assets.[18]

Following the theoretical model above, segment growth opportunity g_i is a parameter about the external environment in which a segment operates. In empirical analysis, we treat g as an industry-level attribute. We identify all single-segment firms in the industry-segment database that operate in the same three-digit SIC industry, and proxy growth opportunity (of an industry) by the *median* consensus long-term growth rate of these single-segment firms forecasted by analysts (on I/B/E/S).[19] Thus, segments with the same three-digit SIC are considered to have the same available growth opportunity (in any given year).[20]

We delete observations with negative book value and trim the top and bottom 0.5 percent extreme values of E/B and V/B.[21] This results in a total of 13,463 observations. Firms having two, three, four, and five or more segments account for 42 percent, 30 percent, 16 percent, and 12 percent of the overall sample, respectively, and the mean (median) number of segments is 3 (3).

Table 1 provides descriptive statistics for scaled market value v, firm profitability q, and *DOP*, both for individual years and for the pooled sample. The number of observations

---

[17] Given that segment-level equity book values are not reported for the period covered in our sample (under SFAS No. 14), segment-level returns on equity cannot be directly computed. As a result, return on operating assets is used to measure segment profitability for the purpose of calculating the divergence of profitability (*DOP*). An alternative measure of *DOP* can be constructed based on imputed segment-level returns on equity whereby total liabilities and interest expenses of a firm are allocated to segments in proportion to segment assets. The two proxies of *DOP* are highly correlated (Pearson correlation coefficient = 0.78). Qualitatively similar empirical results are obtained for the two proxies.

[18] In this more general version, the definition of *DOP* takes into account (1) the "distance" between each segment's profitability and the weighted average profitability of the firm and (2) the asset sizes of segments. Note that for a two-segment firm and given the segments' assets, *DOP* as defined in Equation (4) is equivalently represented by $q_2 - q_1$ (as in the model in Section III). In calculating the weights for empirical analysis, we use the sum of identifiable assets of the firm's segments as denominator, rather than total assets of the firm, because in some cases identifiable segment assets do not add up to the total assets of the firm (this is also noted in Berger and Ofek [1995]).

[19] We need to point out that interpretation of growth parameter g is somewhat different between the model and the empirical setting. In the model, a growth option represents a one-time only growth for a business; therefore, the size of growth option g should be viewed as the "total growth" that can be achieved over time. Empirically, on the other hand, g is proxied by the forecasted *annual* earnings growth rate, and this annual value normally would be much lower than total (cumulative) growth potential.

[20] We use single-segment firms, rather than all firms, in a three-digit SIC industry to match more closely between the segment being measured and the reference firms and also to avoid introducing spurious correlation (as single-segment firms are not in our test samples). Analyst forecasts are taken four months after the fiscal year-end to reflect possible revisions after the release of financial statements. Our growth proxy captures growth from positive NPV opportunities to the extent that analysts forecast earnings growth rates based on their projection of what positive NPV projects will be available to firms.

[21] Similar truncation procedures are employed in prior studies (e.g., Burgstahler and Dichev 1997; Collins et al. 1999; Barth et al. 1998).

TABLE 1
Descriptive Statistics

		v			q			*DOP*		
Year	**n**	**Mean**	**Median**	**Std.**	**Mean**	**Median**	**Std.**	**Mean**	**Median**	**Std.**
86	993	3.07	1.87	3.59	0.08	0.11	0.18	0.07	0.05	0.08
87	1,044	2.76	1.69	3.42	0.09	0.12	0.19	0.07	0.05	0.12
88	1,060	2.46	1.65	2.76	0.10	0.13	0.19	0.08	0.05	0.12
89	1,061	2.70	1.70	3.10	0.09	0.12	0.20	0.07	0.04	0.13
90	1,071	2.42	1.40	3.14	0.08	0.10	0.19	0.07	0.04	0.14
91	1,077	3.22	1.76	4.28	0.05	0.08	0.20	0.06	0.04	0.09
92	1,088	3.26	1.98	3.64	0.06	0.10	0.22	0.06	0.04	0.07
93	1,161	3.70	2.30	3.95	0.07	0.10	0.22	0.07	0.04	0.10
94	1,231	3.16	2.04	3.39	0.08	0.11	0.23	0.06	0.04	0.07
95	1,314	3.60	2.19	4.01	0.08	0.12	0.24	0.07	0.04	0.10
96	1,323	3.85	2.37	4.21	0.10	0.12	0.21	0.06	0.04	0.08
97	1,040	3.82	2.53	3.91	0.10	0.13	0.22	0.07	0.03	0.11
86–97 pooled	13,463	2.82	1.95	2.86	0.08	0.11	0.22	0.07	0.04	0.11

v is scaled market value, defined as share price three months after fiscal year-end times the number of shares outstanding and scaled by beginning book value of common equity; and q is firm overall profitability, defined as aggregate earnings available for common stockholders divided by beginning book value of common equity.
$DOP = \Sigma_{i=1}^{k} |q_i - \bar{q}] * w_i$ is divergence of segment profitability of a firm, where q_i is profitability of segment i, defined as segment i's operating earnings divided by its beginning identifiable assets, w_i = weight of segment i equal to segment i's (beginning) identifiable assets divided by the total (beginning) identifiable assets of all reported segments of the firm, $\bar{q} = \Sigma_{i=1}^{k} w_i * q_i$ is weighted average profitability of the firm, and k = number of segments of the firm.

ranges from 993 in year 1986 to 1,323 in year 1996. The mean and median values of v have a generally increasing trend from 1986 to 1997. The mean (median) of q ranges from 0.05 (0.08) to 0.1 (0.13) during this period, and the mean (median) of *DOP* remains steady at around 0.07 (0.04).

V. EMPIRICAL RESULTS

General Effect of Segment Data (H1)

We first test the incremental value effect of *DOP* with broad cross-sectional samples (H1). Following prior empirical studies, two benchmark models are used to control for the effect of aggregate accounting data. The first benchmark expresses scaled equity market value v as a linear function of scaled aggregate earnings (i.e., firm profitability q); see, for example, Burgstahler and Dichev (1997):

$$v = b_1 + b_2 q + \mu. \tag{5}$$

To test H1, we add *DOP* to Equation (5) as an incremental explanatory variable:

$$v = b_1 + b_2 q + b_3 DOP + \mu. \tag{6}$$

The second benchmark is a piece-wise linear model that allows for nonlinearity of equity value in profitability (Burgstahler and Dichev 1997):

$$v = b_1 + b_2D_M + b_3D_H + b_4q + b_5D_Mq + b_6D_Hq + \varepsilon \qquad (7)$$

where D_M and D_H are dummy variables for observations in the middle and high third q-value ranges in the sample, respectively. Again, we add *DOP* to test the incremental effect of segment-level information:

$$v = b_1 + b_2D_M + b_3D_H + b_4q + b_5D_Mq + b_6D_Hq + b_7DOP + \varepsilon. \qquad (8)$$

With segment profitability and growth opportunity positively correlated in our sample,[22] we predict that *DOP* has a positive effect on value.

Table 2 reports the annual and combined regression results for model (5) (Panel A) and model (6) (Panel B). In both models, firm profitability (scaled aggregate earnings) is significantly related to equity value in ten out of the 12 years and for all years combined, consistent with findings of previous studies (e.g., Burgstahler and Dichev 1997; Collins et al. 1999). The results for model (6) show that the coefficient of *DOP* is positive and significant at the 0.01 level in all 12 years individually and combined.

However, the explanatory powers of these linear models are small. For the overall sample, the adjusted R^2 is 1 percent for benchmark model (5), and it increases to 4 percent after adding *DOP* as a second explanatory variable.

Results for models (7) and (8) are reported in Table 3. Consistent with Burgstahler and Dichev (1997), Panel A shows that in model (7), coefficients on profitability q (b_4, b_5, and b_6) are statistically significant in annual regressions and for the combined sample, except for b_5 in three of the 12 years. Panel B shows that *DOP* as an incremental explanatory variable is significant at the 0.01 level in eight of the 12 years and for all years combined, significant at the 0.05 level in two years, and not significant at the 0.05 level in two remaining years.

The explanatory power of piece-wise linear model (7) is 23 percent for the overall sample, substantially higher than that of linear model (5). However, the incremental explanatory power of *DOP* is low, which ranges from 0 to 3 percent for annual samples, and is 1 percent for the overall sample.[23]

Together, the results in Tables 2 and 3 are consistent with H1 that *DOP* generally has a positive impact on firm value, incremental to aggregate accounting data,[24] but the incremental explanatory power of *DOP* is generally low.

Firm-Level Characteristics and the Incremental Value Effect of *DOP* (H2 and H3)

After showing the incremental value effect of *DOP* in broad-based samples, we next explore how the effect varies across firms.

Firm Overall Profitability and the Effect of DOP (H2)

Hypothesis 2 is examined with two different samples. We first use the pooled sample to gain a general view, ignoring the assumption underlying H2 that a firm's segments have

---

[22] For our overall sample, the Pearson correlation between segment profitability and growth opportunity is +0.07, significant at the 1 percent level.

[23] Low incremental explanatory power of segment data is also found (or implied) in prior studies, e.g., Basu et al. (1999) and Givoly et al. (1999), that use different regression specifications. Our analysis ignores various costs arising in multisegment organizations due to coordination and communication, loss of managerial focus, cross-subsidization, etc. Such costs may offset to some extent the value effect of *DOP* and contribute to low incremental R^2s of *DOP*.

[24] Our result that *DOP* has an incremental value effect is consistent with Wysocki's (1998) finding that earnings coefficients differ between profit and loss segments and between high- and low-profit segments.

TABLE 2
Incremental Value Effect of Divergence of Segment Profitability (*DOP*) from Linear Regressions

Panel A: Benchmark Model (5): $v = b_1 + b_2q + \mu$

Year	n	b_1	t_1	b_2	t_2	Adj. R^2
86	993	2.20	27.00	1.99	4.90**	0.02
87	1,044	2.16	22.30	1.49	3.31**	0.01
88	1,060	2.27	24.91	0.21	0.50	0.00
89	1,061	2.07	29.50	2.58	8.52**	0.06
90	1,071	1.80	25.62	2.03	6.19**	0.03
91	1,077	2.19	26.61	1.49	3.89**	0.01
92	1,088	2.50	31.03	0.24	0.67	0.00
93	1,161	2.84	31.20	1.39	3.67**	0.01
94	1,231	2.49	31.62	0.65	1.98*	0.01
95	1,314	2.80	33.53	0.75	2.34**	0.01
96	1,323	2.88	30.59	1.64	4.09**	0.01
97	1,040	3.17	30.68	2.68	6.60**	0.04
86–97 pooled	13,463	2.43	106.38	1.44	14.60**	0.01

Panel B: Model (6): $v = b_1 + b_2q + b_3DOP + u$

Year	n	b_1	t_1	b_2	t_2	b_3	t_3	Adj. R^2
86	993	1.78	16.5	2.47	6.06	5.40	5.93**	0.06
87	1,044	1.70	15.8	1.73	3.96	6.05	8.67**	0.08
88	1,060	1.95	18.3	0.51	1.26	3.94	5.74**	0.03
89	1,061	1.84	23.4	2.73	9.14	3.02	6.25**	0.10
90	1,071	1.63	21.2	2.06	6.34	2.43	5.13**	0.06
91	1,077	1.76	17.4	1.92	5.03	6.37	7.03**	0.06
92	1,088	2.24	20.5	0.44	1.22	4.30	3.63**	0.01
93	1,161	2.41	22.3	1.77	4.74	6.11	7.01**	0.05
94	1,231	2.14	20.6	0.84	2.58	5.72	5.17**	0.02
95	1,314	2.44	24.1	1.17	3.59	4.68	6.08**	0.03
96	1,323	2.53	20.7	2.03	4.97	4.93	4.31**	0.02
97	1,040	2.73	22.4	3.37	8.19	5.66	6.58**	0.08
86–97 pooled	13,463	2.13	79.6	1.80	18.20	3.89	20.80**	0.04

*, ** Denotes significance at the 0.05 (0.01) level, two-tailed.
t_i represents the t-statistic for coefficient b_i.

equal growth opportunities. We then perform a more precise test on a subsample that approximately satisfies the equal growth opportunity assumption.

For the overall sample, we partition observations into quintiles by firm profitability q and run separate regressions for each of the quintiles. Table 4 reports the descriptive statistics (Panel A) and the regression results based on model (6) (Panel B).

Controlling for firm-level profitability q, the coefficient of *DOP* (b_3) is small (0.39) and insignificant for the extreme-low profitability quintile q1 ($q < 0.1$ percent); it then increases to 2.76 in the second quintile q2 (0.1 percent $\leq q < 8.3$ percent) before decreasing to 1.61 for q3 (8.3 percent $\leq q < 13.4$ percent) and further to 0.70 for q4 (13.4 percent

TABLE 3
Incremental Value Effect of Divergence of Segment Profitability (*DOP*) from Piece-wise Linear Regressions

Panel A: Benchmark Model (7): $v = b_1 + b_2D_M + b_3D_H + b_4q + b_5D_M\,q + b_6D_H\,q + \varepsilon$

Year	b_1	t_1	b_2	t_2	b_3	t_3	b_4	t_4	b_5	t_5	b_6	t_6	Adj. R^2
86	1.34	9.20	−0.20	−0.51	−1.44	−4.39	−4.96	−7.86	10.65	3.08	20.29	15.30	0.27
87	1.43	7.88	−0.13	−0.27	−0.92	−2.37	−4.90	−6.43	8.61	1.99	15.70	10.76	0.15
88	1.40	8.01	0.30	0.69	−1.02	−2.64	−5.52	−8.48	5.52	1.51	15.80	11.23	0.15
89	1.49	9.77	0.15	0.41	−0.59	−2.22	−1.19	−2.21	1.99	0.67	9.70	10.29	0.21
90	1.31	9.15	−0.48	−1.49	−1.30	−4.91	−2.91	−5.37	8.44	2.76	12.74	14.85	0.25
91	0.64	4.04	0.43	1.33	−0.56	−1.86	−6.94	−12.13	14.03	4.29	21.92	19.66	0.33
92	1.35	10.60	−0.37	−0.86	−0.57	−2.11	−6.35	−13.91	14.73	3.43	18.62	18.29	0.33
93	1.47	9.00	0.16	0.42	−1.06	−3.07	−5.65	−11.12	10.63	3.20	21.37	17.91	0.32
94	1.72	10.89	−0.28	−0.76	−0.66	−2.06	−4.20	−8.82	7.36	2.29	14.09	13.03	0.21
95	1.99	11.14	−0.40	−1.04	−1.04	−3.13	−3.78	−7.82	7.81	2.32	15.00	13.84	0.20
96	2.19	14.89	−0.63	−0.85	−1.13	−3.25	−4.50	−7.90	9.68	1.66	16.65	13.61	0.19
97	2.16	9.76	−0.26	−0.56	−1.56	−3.93	−3.61	−5.93	7.56	2.00	19.60	15.33	0.29
86–97 pooled	1.59	43.20	−0.40	−2.66	−1.14	−13.3	−4.24	−30.62	10.33	7.84	17.03	54.78	0.23

(continued on next page)

TABLE 3 (continued)

Panel B: Model (8), $v = b_1 + b_2D_M + b_3D_H + b_4q + b_5D_Mq + b_6D_Hq + b_7DOP + \varepsilon$

Year	b_1	t_1	b_2	t_2	b_3	t_3	b_4	t_4	b_5	t_5	b_6	t_6	b_7	t_7	Adj. R^2
86	1.18	7.61	−0.21	−0.54	−1.37	−4.18	−4.57	−7.10	10.60	3.09	19.62	14.59	2.20	2.70**	0.28
87	1.13	6.23	−0.13	−0.26	−0.66	−1.71	−4.16	−5.50	8.41	1.98	13.68	9.42	4.25	6.28**	0.18
88	1.20	6.58	0.30	0.68	−0.93	−2.46	−5.15	−7.86	5.83	1.60	15.13	10.73	2.42	3.73**	0.16
89	1.49	9.77	0.15	0.41	−0.59	−2.22	−1.19	−2.21	−1.99	0.67	9.74	10.33	1.64	3.52**	0.21
90	1.22	8.36	−0.49	−1.53	−1.22	−4.61	−2.80	−5.19	8.72	2.86	14.31	14.39	1.24	2.89**	0.26
91	0.53	3.22	0.42	1.34	−0.49	−1.62	−6.60	−11.20	13.82	4.23	21.21	18.21	2.02	2.66**	0.33
92	1.30	9.24	−0.37	−0.87	−0.55	−2.05	−6.29	−13.60	14.88	3.47	18.48	18.03	0.80	0.81	0.33
93	1.31	7.68	0.16	0.44	−0.96	−2.77	−5.32	−10.40	10.68	3.20	20.61	17.09	2.56	3.40**	0.33
94	1.58	9.27	−0.26	−0.72	−0.63	−2.00	−4.06	−8.46	7.49	2.33	13.83	12.72	2.14	2.12*	0.22
95	1.84	9.91	−0.39	−1.01	−0.99	−2.98	−3.50	−7.08	7.95	2.36	14.48	13.21	2.04	2.92**	0.21
96	2.16	13.42	−0.62	−0.84	−1.10	−3.23	−4.54	−7.50	9.63	1.65	16.58	13.39	0.41	0.38	0.20
97	2.06	9.09	−0.29	−0.62	−1.46	−3.67	−3.27	−5.18	7.79	2.06	18.96	14.32	1.49	1.90*	0.30
86–97 pooled	1.51	39.21	−0.41	−2.72	−1.07	−12.70	−3.97	−28.01	10.43	7.93	16.53	52.31	1.35	7.93**	0.24

---

*, ** Denotes significance at the 0.05 (0.01) level, two-tailed.

D_M and D_H are dummy variables for firms in the middle- and high-third q-value ranges, respectively, in a sample.

t_i represents the t-statistic for coefficient b_i.

TABLE 4
Firm Profitability and the Incremental Value Effect of *DOP*

Panel A: Descriptive Statistics of Partitions by Firm's Overall Profitability

		v			q			*DOP*		
Partitions by Firm Profitability	**n**	**Mean**	**Med.**	**Std. Dev.**	**Mean**	**Med.**	**Std. Dev.**	**Mean**	**Med.**	**Std. Dev.**
q1 ($q < 0.001$)	2,693	2.40	1.14	3.62	−0.22	−0.13	0.26	0.11	0.06	0.15
q2 ($0.001 \leq q < 0.083$)	2,692	1.63	1.22	1.70	0.05	0.05	0.02	0.06	0.04	0.09
q3 ≤ 0.083 ($q < 0.134$)	2,692	1.85	1.56	1.34	0.11	0.11	0.01	0.05	0.03	0.07
q4 ≤ 0.134 ($q < 0.196$)	2,692	2.35	2.06	1.38	0.16	0.16	0.02	0.05	0.03	0.08
q5 $\geq q$ (0.196)	2,693	4.45	3.61	3.22	0.30	0.26	0.12	0.08	0.05	0.10

Panel B: Model (6): $v = b_1 + b_2q + b_3DOP + \mu$

Partitions by Firm Profitability	**n**	b_1	t_1	b_2	t_2	b_3	t_3	**Adj. R^2**
q1 ($q < 0.001$)	2,693	1.38	15.80	−4.57	−17.80	0.39	1.01	0.10
q2 ($0.001 \leq q < 0.083$)	2,692	1.31	18.30	3.05	2.40	2.76	7.98**	0.02
q3 ($0.083 \leq q < 0.134$)	2,692	1.11	5.70	6.08	3.50	1.61	4.46**	0.01
q4 ($0.134 \leq q < 0.196$)	2,692	−0.22	−1.00	15.68	11.80	0.70	2.29*	0.05
q5 ($q \geq 0.196$)	2,693	0.78	5.70	11.40	27.10	2.38	5.66**	0.21

Panel C: Model (6): $v = b_1 + b_2q + b_3DOP + \mu$ for Subsample of Firms with Similar Segment Growth Opportunities ($n = 2{,}346$)

Partitions by Firm Profitability	**n**	b_1	t_1	b_2	t_2	b_3	t_3	**Adj. R^2**	**Incremental R^2**
q1 ($q < 0.001$)	242	1.46	13.10	−1.02	−1.52	0.27	0.19	0.002	−0.003
q2 ($0.001 \leq q < 0.083$)	341	1.09	7.20	3.84	2.49	3.85	5.25**	0.05	0.045
q3 ($0.083 \leq q < 0.134$)	682	1.07	3.20	5.52	2.27	2.31	3.27**	0.03	0.015
q4 ($0.134 \leq q < 0.196$)	700	−0.69	−1.25	16.99	5.73	3.01	2.67**	0.09	0.022
q5 $\geq q$ (0.196)	381	1.08	2.44	11.39	8.66	1.31	0.65	0.19	−0.002

---

*, ** Denotes significance at the 0.05 (0.01) level, two-tailed.
t_i represents the t-statistic for coefficient b_i.
Incremental R^2 is the adjusted R^2 of model (6) minus the adjusted R^2 without *DOP* in model (6).

$\leq q <$ 19.6 percent). The coefficient then rises to 2.38 for the extreme-high profitability quintile q5 ($q \geq$ 19.6 percent). Coefficient b_3 is significant for quintiles q2 to q5.

These results are generally consistent with H2 in low-profitability ranges, which predicts that *DOP* has little incremental effect for extreme-low q-values, and, as q increases, its effect first rises and then falls. However, the relatively large *DOP* coefficient in the highest profitability quintile q5 is at odds with H2 where the incremental effect is predicted to be small. This discrepancy arises possibly because the assumption underlying H2 (and Figure 2), that a firm's segments have equal growth opportunities, is not always satisfied empirically. Note that this assumption should have little consequence in low-profitability regions where growth options are deep out of the money.

To test H2 more precisely, we partition the overall sample into thirds by difference in segment growth opportunity Δg, and perform analysis only on the middle third where segments within a firm have similar growth opportunities (i.e., Δg is close to zero).[25] The results are presented in Panel C of Table 4. The coefficients of *DOP* (b_3) based on this subsample are 0.27, 3.85, 2.31, 3.01, and 1.31 respectively for profitability partitions q1 to q5, insignificant for both the lowest and the highest partitions (q1 and q5) but significant for the intermediate ones. The result for low-profitability quintiles is qualitatively the same as from the broad sample before (Panel B), which is consistent with H2, and the result for high-profitability quintiles is in contrast to that in Panel B and is now also consistent with H2. Moving across the entire profitability range from q1 to q5, the pattern of variation of b_3 is similar to what is predicted by H2 (and Figure 2).

Panel C also reports the incremental R^2s of *DOP*. The incremental R^2 of *DOP* is virtually zero for both the lowest and the highest profitability quintiles, q1 and q5; it obtains a higher value of 4.5 percent for q2, 1.5 percent for q3, and 2.2 percent for q4. While the incremental explanatory power of *DOP* is still generally small, it shows an obvious variation across different profitability ranges; the pattern of variation is the same as that of the *DOP* coefficient, and it is consistent with H2.

Firm Overall Growth and the Effect of DOP (H3)

To explore how the value effect of segment data changes with firm overall growth opportunity g (H3), we partition the pooled sample into quintiles by g, whereby firm overall growth is proxied by analysts' consensus forecast of a firm's long-term growth rate. The availability of such forecasts reduces the sample size to 7,040. Table 5 reports the descriptive statistics (Panel A) and results (Panels B and C). Results based on linear model (6), in Panel B, shows that the coefficient of *DOP*, b_3, exhibits a monotonically increasing trend with firm overall growth opportunity g across all growth-quintiles, except for the highest growth quintile g5. The increases in *DOP* coefficient (b_3) for higher growth partitions, relative to lowest partition g1, are 1.15, 2.54, 4.60, and 1.75 for partitions g2 to g5, with corresponding t-statistics of 1.61 ($p = 0.06$), 2.54 ($p = 0.002$), 5.03 ($p = 0.000$), and 1.17 ($p = 0.12$), respectively, showing an increasing trend in b_3 except for g5.

Panel C provides the results for the piece-wise linear model (8). The coefficient of *DOP*, b_7, exhibits the same general pattern across growth partitions, as in Panel B. The increases in b_7 for higher growth partitions relative to g1 are now 0.45, 1.45, 1.82 and −0.62 for partitions g2 to g5, with corresponding t-statistics of 0.64 ($p = 0.26$), 1.82 ($p = 0.04$), 2.12 ($p = 0.02$) and −0.52 ($p = 0.70$), respectively, again showing a generally increasing trend of b_3 except for g5.[26]

Thus, based on both the linear and the piece-wise linear model, except for highest growth quintile g5, the magnitude of the *DOP* coefficient exhibits an increasing trend, consistent with H3 that the incremental value effect of *DOP* increases with firm overall growth opportunity. On the other hand, the incremental R^2 of *DOP* (reported in Panel C for the piece-wise linear model) remains small, ranging from 0 to 1 percent.

---

[25] For comparison, we apply the same set of cut-off values of profitability for partitioning this subsample as in Panel B of Table 4. The empirical definition of Δg is provided below, which generalizes that in the theoretical model. Also note that requiring analyst forecasts of earnings growth as growth proxies reduces the overall sample size to 7,040.

[26] The smaller coefficient of *DOP* for the highest firm-level growth partition could be due to the possibility that growth opportunities are more concentrated in relatively low-profitability segments for firms within this partition compared with other partitions. As will be shown below, the distribution of growth opportunities within a firm also matters.

TABLE 5
Firm Growth Opportunity and the Incremental Value Effect of *DOP*

Panel A: Descriptive Statistics of Partitions by Analysts' Forecast of Firm's Long-Term Growth Rate

Partitions by Firm Growth Opportunity	n	v			q			*DOP*			g(%)		
		Mean	Med.	Std. Dev.	Mean	Med.	Std. Dev.	Mean	Med.	Std. Dev.	Mean	Med.	Std. Dev.
$g1$ ($g < 8.0$)	1,408	1.57	1.47	0.72	0.10	0.12	0.12	0.03	0.02	0.04	5.37	5.50	2.08
$g2$ ($8.0 \leq g < 10.0$)	1,408	1.94	1.66	1.18	0.12	0.12	0.12	0.05	0.04	0.05	9.60	10.00	0.59
$g3$ ($10.0 \leq g < 12.0$)	1,408	2.48	2.02	1.74	0.14	0.15	0.12	0.05	0.04	0.05	11.68	12.00	0.64
$g4$ ($12.0 \leq g < 15.0$)	1,408	2.89	2.36	1.89	0.16	0.16	0.13	0.06	0.04	0.06	14.35	15.00	0.77
$g5$ ($g \geq 15.0$)	1,408	3.79	2.96	3.05	0.15	0.15	0.16	0.06	0.04	0.06	20.92	20.00	4.97

Panel B: Model (6): $v = b_1 + b_2q + b_3DOP + \mu$

Partitions by Firm Growth Opportunity (%)	n	b_1	t_1	b_2	t_2	b_3	t_3	Adj. R^2
$g1$ ($g < 8.0$)	1,408	1.25	44.27	2.59	17.69	1.87	4.16**	0.18
$g2$ ($8.0 \leq g < 10.0$)	1,408	1.22	27.67	5.43	25.17	1.90	3.60**	0.32
$g3$ ($10.0 \leq g < 12.0$)	1,408	1.12	16.01	8.33	26.01	3.19	3.97**	0.34
$g4$ ($12.0 \leq g < 15.0$)	1,408	1.26	17.90	8.57	28.37	4.91	7.34**	0.39
$g5$ ($g \geq 15.0$)	1,408	2.45	20.67	8.34	18.59	2.30	2.04*	0.20

(*continued on next page*)

TABLE 5 (continued)

Panel C: Model (8): $v = b_1 + b_2D_M + b_3D_H + b_4q + b_5D_Mq + b_6D_Hq + b_7DOP + \varepsilon$

Partitions by Firm Growth Opportunity (%)	b_1	t_1	b_2	t_2	b_3	t_3	b_4	t_4	b_5	t_5	b_6	t_6	b_7	t_7	Adj. R^2	Incremental R^2
$g1$ ($g < 8.0$)	1.15	34.76	−0.46	−1.71	0.13	1.63	0.28	1.25	6.82	2.93	2.81	6.11	0.98	2.26*	0.21	0.003
$g2$ ($8.0 \leq g < 10.0$)	1.29	27.13	−0.52	−2.14	−0.73	−6.13	1.40	3.29	5.43	2.79	7.89	12.76	1.43	2.75**	0.47	0.004
$g3$ ($10.0 \leq g < 12.0$)	1.61	27.74	−1.65	−4.55	−1.99	−14.36	−0.85	−1.52	14.92	5.84	16.01	21.59	2.44	3.92**	0.57	0.005
$g4$ ($12.0 \leq g < 15.0$)	1.67	23.78	−1.32	−2.89	−1.78	−7.97	−0.01	−0.02	12.55	4.31	15.54	16.29	2.80	4.63**	0.51	0.010
$g5$ ($g \geq 15.0$)	2.41	19.51	−0.87	−1.41	−0.66	−2.07	−1.02	−1.16	10.86	2.68	14.19	11.14	0.22	0.21	0.36	0.000

---

*, ** Denotes significance at the 0.05 (0.01) level, two-tailed.

D_M and D_H are dummy variables for firms in the middle and high-third q-value ranges, respectively, in a sample.

t_i represents the t-statistic for coefficient b_i.

Incremental R^2 is the adjusted R^2 of model (8) minus the adjusted R^2 without *DOP* in model (8).

Segment Growth Opportunities and the Value Effect of *DOP* (H4)

Our model further predicts that the incremental value effect of *DOP* depends on the distribution of growth opportunities within the firm. We now conduct tests that explicitly recognize heterogeneity of growth opportunities across segments (H4).

Testing H4 requires a measure of the extent to which more profitable segments have more growth opportunity than less profitable segments (Δg). For a two-segment firm considered in the theoretical model, Δg is defined as follows: label the less profitable segment as segment 1 and the other as segment 2, and let g_1 and g_2 be their respective growth opportunities. The heterogeneity in growth opportunities between segment 2 and segment 1, adjusted for segment size differences, is defined as:

$$\Delta g = w_1(\bar{g} - g_1) + w_2(g_2 - \bar{g}) \tag{9}$$

where $\bar{g} = g_1 * w_1 + g_2 * w_2$ is the weighted average growth opportunity of the firm, with weights w_1 and w_2 defined in proportion to beginning book values of identifiable assets.[27]

For general multiple-segment firms, we generalize Equation (9) as follows:

$$\Delta g = \sum_i w_i(\bar{g} - g_i) + \sum_j w_j(g_j - \bar{g}) \tag{10}$$

where $\bar{g}$ is the weighted average growth opportunity; segment *i*'s (first summation sign) correspond to the relatively unprofitable segments with profitability below the weighted average profitability $\bar{q}$ (defined in Section IV), and *j*'s (second summation sign) correspond to relatively profitable segments with profitability above $\bar{q}$.

Our model predicts that both the sign and the magnitude of Δg are important in determining the value effect of *DOP*. A positive Δg indicates that most of the firm's growth opportunities lie in relatively profitable segments, whereas a negative Δg indicates the opposite.

Following the theoretical analysis, H4 applies to firms with high profitability. We divide the overall sample into the low- and high-profitability halves. Testing of H4 is conducted only on the high-profitability half, which is further partitioned equally into the low-, middle-, and high-Δg groups. We examine how the value effect of *DOP* changes across firms with different Δg-values.[28]

Panel A of Table 6 reports descriptive statistics for three Δg-partitions. Panel B reports the test results. Both the sign and the magnitude of the coefficient of *DOP* (b_3) are found to relate closely to those of Δg in a way predicted by H4. Specifically, in the low-Δg partition (where $\Delta g < 0$ for all observations), *DOP* has a negative incremental effect on firm value; coefficient b_3 is estimated to be −1.60, significant at the 5 percent level for one-tailed tests.[29] In the middle-Δg partition, where more than half of the observations have a positive-Δg (median $\Delta g = 0.12$), the value effect of *DOP* is positive, with $b_3 = 1.83$, significant at the 5 percent level. The high-Δg partition consists purely of positive-Δg observations, and the effect of *DOP* is positive, with $b_3 = 3.23$, significant at the 1 percent

---

[27] The definition of Δg here takes into account both the "distance" and the "direction" of each segment's growth opportunity relative to the average growth opportunity as weighed by segment assets. It can be shown that this definition is equivalent to $[A_1A_2/(A_1 + A_2)]^2(g_2 - g_1)$, an expression that arises from analytical derivations to capture the effect of the divergence in segment growth opportunity.

[28] Results in Table 6 (and also Table 7 below) are based on regression models (5) and (6). Similar results obtain if models (7) and (8) are used.

[29] We apply one-tailed tests for H4 because the hypothesis predicts the direction (sign) of the value effect of *DOP*.

TABLE 6
Divergence of Segment Growth Opportunity (Δg) and the Incremental Value Effect of *DOP*

Panel A: Descriptive Statistics of Partitions by Divergence of Segment Growth Opportunity

Partitions		v			q			*DOP*			Δg(%)		
by Growth Divergence	**n**	**Mean**	**Med.**	**Std. Dev.**	**Mean**	**Med.**	**Std. Dev.**	**Mean**	**Med.**	**Std. Dev.**	**Mean**	**Med.**	**Std. Dev.**
Low Δg	1,173	3.44	2.82	2.32	0.22	0.20	0.09	0.05	0.04	0.06	−6.06	−5.70	2.65
Middle Δg	1,173	3.04	2.47	2.05	0.21	0.18	0.09	0.06	0.04	0.06	0.00	0.12	1.30
High Δg	1,174	3.46	2.82	2.36	0.23	0.19	0.10	0.05	0.04	0.06	5.36	4.91	2.30

Panel B: Regression Results on the Effect of DOP on Equity Value

	b_1	t_1	b_2	t_2	b_3	t_3	Adj. R²
Model (5): $v = b_1 + b_2q + u$							
Low Δg	−0.31	−2.34	16.79	30.54			0.45
Middle Δg	0.06	0.49	14.19	27.49			0.39
High Δg	0.01	0.09	15.22	29.06			0.41
Model (6): $v = b_1 + b_2q + b_3DOP + u$							
Low Δg	−0.26	−1.90	16.94	30.48	−1.60	−1.81*	0.45
Middle Δg	0.03	0.29	13.81	25.34	1.83	2.12*	0.40
High Δg	−0.08	−0.59	14.86	27.99	3.23	3.47**	0.42

*, ** Denotes significance at the 0.05 (0.01) level, one-tailed.
t_i represents the t-statistic for coefficient b_i.
This table reports the results on the value effect of the difference in growth opportunities between segments (Δg) based on firms with high overall profitability. The data are partitioned into three equal-sized partitions by Δg.
$\Delta g = \Sigma_i w_i(\bar{g} - g_i) + \Sigma_j w_j(g_j - \bar{g})$ is the difference in growth opportunities between segments, where g_i is the growth opportunity of a segment with profitability less than firm average and g_j is the growth opportunity of a segment with profitability greater than firm average; $\bar{g}$ is the weighted average growth opportunities of a firm; and w_i and w_j are weights defined in terms of segment identifiable assets.

level. The estimates of b_3 for the three Δg partitions also confirm that the incremental value effect of *DOP* increases monotonically with Δg. However, the incremental R² of *DOP* is small, 0 percent for low-Δg partition and 1 percent for middle- and high-Δg partition.

We have shown previously that firm overall growth opportunity (g) affects the value effect of *DOP*. To gain additional insight, we examine whether the result concerning H4 differs for firms with different overall growth opportunities. We divide the above sample employed for H4 equally into two groups, low-g and high-g, and repeat the test for each group. Panel A of Table 7 reports the descriptive statistics, and Panels B and C report regression results for the low- and the high-g group, respectively.

For both the low-g (Panel B) and the high-g group (Panel C), the result exhibits the same general pattern as for the pooled sample (Table 6). That is, as Δg increases from negative to positive values, the value effect of *DOP* also increases from negative to positive values. Thus, the sign and the magnitude of coefficient b_3 relate to those of Δg in a way as predicted by the model (H4).[30] However, the incremental R² of *DOP* remains small as in Table 6.

---

[30] Relatedly, Tse (1989) and Wysocki (1998) show that differences in growth opportunity among segments affect the usefulness of segment data. An insight from our study is that segment growth opportunities interact with segment profitability in determining value, and one needs to distinguish the effect of Δg between high- and low-profitability region.

TABLE 7
Divergence of Segment Growth Opportunity (Δg) and the Incremental Value Effect of *DOP* Controlling for Firm Growth Opportunity (g)

Panel A: Descriptive Statistics by Partitions

Partitions by Growth Divergence	n	v Mean	v Med.	v Std. Dev.	q Mean	q Med.	q Std. Dev.	*DOP* Mean	*DOP* Med.	*DOP* Std. Dev.	Δg(%) Mean	Δg(%) Med.	Δg(%) Std. Dev.
Low firm overall growth opportunity													
Low Δg	554	2.56	2.17	1.53	0.20	0.18	0.08	0.05	0.04	0.05	−6.01	−5.76	2.55
Middle Δg	675	2.24	1.95	1.13	0.19	0.17	0.08	0.05	0.03	0.05	0.05	0.19	1.27
High Δg	530	2.53	2.25	1.30	0.21	0.19	0.08	0.05	0.03	0.05	4.91	4.58	1.98
High firm overall growth opportunity													
Low Δg	618	4.04	3.42	2.56	0.24	0.21	0.10	0.06	0.04	0.06	−6.10	−5.65	2.72
Middle Δg	499	3.84	3.20	2.41	0.23	0.20	0.09	0.06	0.05	0.06	−0.05	−0.03	1.33
High Δg	644	3.99	3.25	2.65	0.24	0.20	0.11	0.06	0.04	0.06	5.62	5.14	2.43

Panel B: Regression Results for the Low Firm Overall Growth Opportunity Partition

	b_1	t_1	b_2	t_2	b_3	t_3	Adj. R^2
Model (5): $v = b_1 + b_2q + u$							
Low Δg	−0.18	−1.37	13.48	22.30			0.51
Middle Δg	0.71	7.30	7.94	16.98			0.33
High Δg	0.45	3.38	9.98	16.69			0.39
Model (6): $v = b_1 + b_2q + b_3DOP + u$							
Low Δg	−0.15	−1.12	13.54	22.18	−0.82	−0.77	0.51
Middle Δg	0.71	7.28	7.74	15.54	0.88	1.17	0.33
High Δg	0.38	2.87	9.74	16.00	2.55	2.35**	0.39

(*continued on next page*)

TABLE 7 (continued)

Panel C: Regression Results for the High Firm Overall Growth Opportunity Partition

	b_1	t_1	b_2	t_2	b_3	t_3	Adj. R^2
Model (5): $v = b_1 + b_2q + u$							
Low Δg	−0.01	−0.01	17.06	22.05			0.41
Middle Δg	0.04	0.21	16.60	20.86			0.43
High Δg	0.19	1.10	16.02	23.56			0.42
Model (6): $v = b_1 + b_2q + b_3DOP + u$							
Low Δg	0.09	0.43	17.31	22.19	−2.53	−2.15*	0.42
Middle Δg	0.00	0.01	16.36	19.80	1.46	1.08	0.43
High Δg	0.11	0.62	15.74	22.82	2.63	2.20**	0.43

See the explanatory notes for Table 6.

A further observation is that the extent of change of b_3 across different Δg partitions differs between the low- and the high-g group. The changes are more sensitive for firms with high growth opportunity. Among firms with low overall growth opportunity, coefficient b_3 increases from -0.82 for the low-Δg partition, to 0.88 for the medium-Δg partition, and to 2.55 for the high-Δg partition; the change of b_3 from low- to high-Δg partition is 3.37. On the other hand, among firms with high growth opportunity, the corresponding b_3 estimates are -2.53, 1.46, and 2.63; the change of b_3 from low- to high-Δg partition is 5.16.

Further Analysis for H4

The above empirical analysis shows that the observed variations in regression coefficients are consistent with the model's predictions, but the incremental explanatory power is generally small, particularly for results concerning H4. One possible cause for low incremental explanatory power of *DOP*, which is much lower than the explanatory power of firm-level profitability q, is that reported segment data might contain considerable measurement errors. The quality of segment data has been a concern for users of financial information (see, for example, AIMR 1993; AICPA 1994). Givoly et al. (1999) present evidence that segment data are less precise relative to firm-level data.

Relative to firm-level data, segment data are subject to more measurement errors due to common cost allocation, inter-segment transfer pricing, and imprecise identification of a segment's industry, etc. Measurement error problems are likely to be more severe for firms that are organizationally more complex and their segments more operationally interdependent (Givoly et al. 1999). In an attempt to reduce the effect of measurement errors, we select a subsample of two-segment firms (less complex organizational structure) with segments operating in different two-digit SIC industries (dissimilar growth opportunities), to further test H4. The procedure for Table 6 is repeated, except that now we partition observations into quintiles by Δg (hoping to amplify the cross-sectional variations).

Results reported in Table 8 based on this subsample show that the coefficient of *DOP*, b_3, increases monotonically with the difference in segment growth opportunity Δg; the sign and the magnitude of b_3 are closely aligned with those of Δg. While qualitatively similar to the results from broad samples above (Tables 6 and 7), the *DOP* coefficient now becomes more significant for partitions with large absolute values of Δg, both negative and positive, where the model predicts that the effect should be the greatest, exhibiting a more clear variation across different partitions of Δg. Furthermore, there is marked improvement in the incremental explanatory power of *DOP*. The incremental R^2s for the five Δg partitions are 2.2 percent, 0 percent, 0.1 percent, 3.3 percent, and 5.4 percent, respectively, exhibiting a trend closely related to the (absolute) magnitude of Δg. These improved regression results provide stronger evidence consistent with H4 than those from broader samples above.

An Assessment of the Economic Significance of Segment Data

We now assess the economic significance of segment information by addressing the issue of *how much* segment information matters in valuation. We estimate the extent to which (average) equity value changes with segment information, based on the regression results. We then explain why the demonstrated economic significance (as will be shown below) does not contradict the generally low incremental explanatory power reported.

To illustrate the importance of segment information, take as an example the estimated valuation function for range 0.1 percent $< q <$ 8.3 percent, shown in Panel B, Table 4, $v = 1.31 + 3.05q + 2.76DOP$. At the mean value of $q = 0.05$, $v = 1.4625 + 2.76DOP$. Based on this estimated equation, a change of *DOP* by 0.05 will, on average, result in a change of firm value by approximately 9 percent, and a change of *DOP* by 0.10 will result

TABLE 8
Divergence of Segment Growth Opportunity (Δg) and the Incremental Value Effect of *DOP*
A Subsample of Two-Segment Firms with Operations in Two Different Two-Digit SIC Industries

A: Descriptive Statistics of Five Partitions by Δg (885 Observations with High Overall Profitability)

Partitions by Growth Divergence	n	v Mean	v Med.	v Std. Dev.	q Mean	q Med.	q Std. Dev.	*DOP* Mean	*DOP* Med.	*DOP* Std. Dev.	Δg(%) Mean	Δg(%) Med.	Δg(%) Std. Dev.
Low $\Delta g1$	177	3.81	3.19	2.21	0.23	0.20	0.10	0.03	0.02	0.03	−6.86	−6.25	2.25
$\Delta g2$	177	3.11	2.65	1.85	0.21	0.19	0.08	0.04	0.02	0.04	−2.04	−1.95	1.01
$\Delta g3$	177	3.14	2.22	2.12	0.22	0.18	0.11	0.05	0.03	0.05	0.80	0.85	0.63
$\Delta g4$	177	3.29	2.91	1.75	0.23	0.20	0.10	0.05	0.03	0.05	3.30	3.28	0.77
High $\Delta g5$	177	4.02	3.15	2.40	0.25	0.21	0.12	0.04	0.02	0.05	6.77	6.16	1.83

B: Regression Results on the Effect of DOP on Equity Value

	b_1	t_1	b_2	t_2	b_3	t_3	Adj. R^2
Model (5): $v = b_1 + b_2q + u$							
Low $\Delta g1$	0.30	1.22	14.99	15.64			0.473
$\Delta g2$	−0.12	−0.49	15.28	14.01			0.442
$\Delta g3$	0.30	1.39	12.89	15.12			0.471
$\Delta g4$	0.74	3.46	11.17	13.01			0.386
High $\Delta g5$	0.85	3.40	12.42	13.96			0.416
Model (6): $v = b_1 + b_2q + b_3DOP + u$							
Low $\Delta g1$	0.49	1.90	15.06	15.83	−7.15	−2.32**	0.495
$\Delta g2$	−0.10	−0.36	15.28	14.03	−0.76	−0.30	0.442
$\Delta g3$	0.28	1.32	12.48	13.49	2.26	1.14	0.472
$\Delta g4$	0.57	2.67	10.85	12.80	5.25	3.25**	0.419
High $\Delta g5$	0.71	2.89	11.52	13.12	9.28	4.68**	0.470

---

*, ** Denotes significance at the 0.05 (0.01) level, one-tailed.

t_i represents the t-statistic for coefficient b_i.

This table reports the result on the value effect of the difference in growth opportunities between segments (Δg) based on two-segment firms with two segments operating in different two-digit SIC industries.

Δg is defined by Equation (9).

in firm value to change by 18 percent (the standard deviation of *DOP* within this partition is 0.09). Repeating this analysis for other quintiles in Panel B of Table 4 yields the following estimated value changes:[31]

Profitability Range	Change of Equity Value Attributed to DOP Change of 0.05
q1 ($q < 0.001$)	1.0%
q2 ($0.001 \leq q < 0.083$)	9.0%
q3 ($0.083 \leq q < 0.134$)	4.5%
q4 ($0.134 \leq q < 0.196$)	1.4%
q5 $\leq q \geq 0.196$)	3.3%

While the incremental effect of *DOP*, as shown by above percentage value changes, is small *relative to* the effect of firm-level information (q), these percentages are nonetheless significant economically. They suggest that valuation errors from ignoring or inappropriately using segment data can be considerable and can have important consequences for both investors and firms. Furthermore, significant variations exist in the valuation impact of segment data across firms. For example, the effect of *DOP* is negligible for firms in the extreme low-profitability range, but it becomes more significant in higher profitability ranges.

The empirical analysis above also reveal that the value effect of *DOP* can be either positive or negative, depending on the distribution of growth opportunities within the firm (as characterized by Δg). As further illustration of cross-sectional differences, we repeat the above exercise for the regressions reported in Panel C of Table 7, with the following result:

Segmental Growth Differences	Change of Equity Value Attributed to DOP Change of 0.05
Low Δg	−3.0%
Medium Δg	1.9%
High Δg	3.3%

Again, the estimated percentage changes in equity value indicate that the extent of valuation errors can be considerable if one ignores or misuses segment data. Furthermore, the relative importance of segment data varies considerably across firms that face different bundles of investment opportunities. For example, (among firms having high growth opportunity) an increase of *DOP* by 0.05 is on average predicted to result in a 3 percent drop in firm value in the low-Δg group, but a 3.3 percent rise in firm value in the high-Δg group, with a total difference of 6.3 percent between these two groups. (The standard deviation of *DOP* in each partition is about 0.06.) The above estimated equity value changes are generally in line with the magnitudes indicated by simulation results (Figures 2 and 4) in Section III.

While we have shown that on average segment information has a material impact on valuation, the incremental R^2s of *DOP* are generally low. To reconcile these two aspects,

---

[31] The estimations here are for illustration only. Our general conclusion does not change if other reported regression equations are used to compute value changes.

we need to distinguish the two types of information they convey. The value changes estimated from regression results point to an average trend in how sensitively equity value changes with segment information; this is determined by the magnitude and significance of regression coefficients. On the other hand, the incremental R^2 measures the extent to which "residual" terms have been reduced by adding segment information. Loosely speaking, the former corresponds to the mean and the latter to the variance. And it is conceivable that an independent variable obtains a significant coefficient but a low explanatory power.[32] This would happen if, for example, the independent variable, which is theoretically important in explaining the dependent variable, contains large measurement errors. Our evidence justifies what users have expressed, that segment data are important for valuation but the quality of the reported segment data needs to be improved.

VI. CONCLUSIONS AND IMPLICATIONS

In this paper, we apply a real-options-based valuation framework to predict and test the valuation relevance of segment accounting data, incremental to the effect of firm-level accounting data. The paper recognizes that different segments of a firm generally face different investment opportunities and that an important role of segment accounting data is to guide segment-level investment decisions. The key determinants of investment decisions are segment-level profitability and growth opportunities. Segment-level data are incrementally useful because they convey the divergences of profitability and, thus, investment decisions among segments. Our model predicts that the incremental valuation impact of segment data varies systematically with firm overall profitability, overall growth opportunity, and the distribution of growth opportunities within the firm. The empirical results are generally consistent with theoretical predictions.

A general insight of the study is that, given firm-level accounting data, the incremental usefulness of segment data relates to the heterogeneity of investment activities across segments (in terms of whether operations are to expand or to contract). The incremental value relevance of segment data is greater when, for example, some segments are expected to experience growth while others are expected to experience downsizing. On the other hand, disaggregation is not value-relevant for segments that have similar investment prospects.

Our study is useful to users of financial statements. It shows how value-relevant information can be extracted from raw segment accounting data and integrated in valuation, along with firm-level information, and points to circumstances where segment data matter most. The results indicate that, on average, segment data have a material incremental impact on valuation and, thus, ignoring or misusing such data can cause considerable valuation errors. Furthermore, the incremental impact of segment data varies significantly across different firms. On the other hand, the study finds that the incremental explanatory power of segment information is generally low, possibly because reported segment data contain considerable measurement errors, as suggested by Givoly et al. (1999) and our own analysis.[33]

Given the link between internal management decisions and the value relevance of segment data, it is beneficial to organize external reporting of segment data in a way consistent

---

[32] Our results, however, suggest that the estimated economic magnitude attributed to segment information probably has low precision.

[33] Our study does not consider the costs of operating dissimilar segments in a firm due to, for example, communication and coordination, loss of management focus, and cross-subsidization. Such costs can potentially offset some of the predicted value effect of *DOP*, and this might also contribute to the low incremental explanatory power.

with the internal use of such data by management. From external users' perspectives, valuation can be facilitated if firms disclose both the scopes of activities that managerial decisions cover and the types of information used in making decisions. According to our analysis, information pertaining to operating profitability and growth opportunities at the segment level is particularly relevant to valuation. In this sense, our model supports the management approach adopted in SFAS No. 131 and SEC's MD&A requirement to portray the firm's operations "through the eyes of management." Furthermore, our result that segments can be aggregated when they have similar profitabilities and growth opportunities is consistent with the aggregation criteria recommended in SFAS No. 131. However, with generally low explanatory power of segment information, a question posed to regulators and standard setters is: "What measures can they take to reduce the measurement error in segment data?"

For academic research, this study provides insights into how segment-level data could be transformed and incorporated in equity valuation and how empirical data could be partitioned so as to uncover more effectively the valuation impact of segment data and its cross-sectional differences. It will be useful for future research to examine attributes of segment disclosures that can mitigate measurement errors and increase the power of segment information in explaining equity values.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 1994. *Improving Business Reporting—A Customer Focus: A Comprehensive Report of the Special Committee on Financial Reporting.* New York, NY: AICPA.

Association for Investment and Management Research (AIMR). 1993. *Financial Reporting in the 1990s and Beyond.* Charlottesville, VA: AIMR.

Barth, M., W. Beaver, and W. Landsman. 1998. Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting and Economics* 25: 1–34.

Basu, S., O. Kim, and S. Lim. 1999. The usefulness of industry segment information. Working paper, Baruch College–City University of New York.

Berger, P., and E. Ofek. 1995. Diversification's effect on firm value. *Journal of Financial Economics* 37: 39–65.

———, ———, and I. Swary. 1996. Investor valuation of the abandonment option. *Journal of Financial Economics* 42: 257–287.

Biddle, G., P. Chen, and G. Zhang. 2001. When capital follows profitability: Non-linear residual income dynamics. *Review of Accounting Studies* 6: 229–265.

Burgstahler, D., and I. Dichev. 1997. Earnings, adaptation, and equity value. *The Accounting Review* 72 (2): 187–215.

Collins, D., M. Pincus, and H. Xie. 1999. Equity valuation and negative earnings: The role of book value of equity. *The Accounting Review* 74 (January): 29–61.

de Matos, J. 2001. *Theoretical Foundations of Corporate Finance.* Princeton, NJ: Princeton University Press.

Feltham, G., and J. Ohlson. 1995. Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research* 11: 689–731.

———, and ———. 1996. Uncertainty resolution and theory of depreciation measurement. *Journal of Accounting Research* 34 (Autumn): 209–234.

Financial Accounting Standards Board (FASB). 1997. *Disclosures about Segments of an Enterprise and Related Information.* Statement of Financial Accounting Standards No. 131. Stamford, CT: FASB.

Foster, G. 1975. Security price revaluation implications of sub-earnings disclosure. *Journal of Accounting Research* 13 (Autumn): 283–292.

Givoly, D., C. Hayn, and J. D'Souza. 1999. Measurement errors and the information content of segment reporting. *Review of Accounting Studies* 4: 15–43.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20: 125–153.

Hull, J. 2000. *Options, Futures, and Other Derivative Securities*. 4th edition. Upper Saddle River, NJ: Prentice Hall International.

International Accounting Standards Committee (IASC). 1997. *International Accounting Standard IAS 14 (Revised)—Segment Reporting*. London, UK: IASC.

Kester, W. 1984. Today's options for tomorrow's growth. *Harvard Business Review* (March-April): 153–160.

Lobo, G., S. Kwon, and G. Ndubizu. 1998. The impact of SFAS No. 14 segment information on price variability and earnings forecast accuracy. *Journal of Business Finance and Accounting* 15: 969–985.

Myers, S. 1977. Determinants of corporate borrowing. *Journal of Financial Economics* 5: 147–175.

Ohlson, J. 1995. Earnings, book value, and dividends in equity valuation. *Contemporary Accounting Research* 11: 661–687.

Pactor, P. 1993. *Reporting Disaggregated Information*. Financial Accounting Series No. 123-A. Stamford, CT: FASB.

Securities and Exchange Commission (SEC). 1998. *Securities and Exchange Commission v. Sony Corporation*, Case No. 1:98CV01935 (D.D.C.). Accounting and Auditing Enforcement Release No. 162. Litigation Release No. 15832. Washington, D.C.: Government Printing Office.

Tse, S. 1989. Attributes of industry, industry segment and firm-specific information in security valuation. *Contemporary Accounting Research* 5: 592–614.

Wysocki, P. 1998. Real options and informativeness of segment disclosures. Working paper, University of Michigan.

Yee, K. 2000. Opportunities knocking: Residual income valuation of the adaptive firm. *Journal of Accounting, Auditing and Finance* 15: 225–266.

Zhang, G. 2000. Accounting information, capital investment decisions, and equity valuation: Theory and empirical implications. *Journal of Accounting Research* 38 (Autumn): 271–295.

THE ACCOUNTING REVIEW
78, No. 2
2003
pp. 429–448

The Effects of Firm-Wide and Office-Level Industry Expertise on Audit Pricing

Andrew Ferguson
University of New South Wales

Jere R. Francis
University of Missouri–Columbia

Donald J. Stokes
University of Technology, Sydney

ABSTRACT: This study examines the role of auditor industry expertise in the pricing of Big 5 audits in Australia. We test if the audit market prices an auditor's firm-wide industry expertise, or alternatively if the audit market only prices office-level expertise in those specific cities where the auditor is the industry leader. We document that there is an average premium of 24 percent associated with industry expertise when the auditor is *both* the city-specific industry leader and one of the top two firms nationally in the industry. However, the top two firms nationally do not earn a premium in cities where they are not city leaders. We further document that national leadership rankings are, in fact, driven by the specific offices where accounting firms are city leaders. Thus, the overall evidence supports that the market perception and pricing of industry expertise in Australia is primarily based on office-level industry leadership in city-specific audit markets.

Keywords: *audit fees; industry expertise; Big 5 accounting firms.*

Data Availability: *Data are publicly available.*

I. INTRODUCTION

This study investigates the effect of industry expertise on audit fees of Big 5 accounting firms and is motivated by recent findings in Ferguson and Stokes (2002), who document a decline in audit fee premia for industry specialists during the 1990s. Specifically, they report evidence that the audit fee premia documented in Craswell et al. (1995) and based on 1987 data no longer existed by 1998. In light of this result, we re-examine the data in Ferguson and Stokes (2002) using a new research design to determine

---

We appreciate the helpful comments of the two referees and comments received on the paper when presented at University of British Columbia, and the 2002 Annual Meetings of the American Accounting Association and the Accounting Association of Australia and New Zealand, and the 2002 International Symposium on Audit Research.

Editor's note: This paper was accepted by E. Michael Bamber, *Ad Hoc* Editor.

Submitted January 2002
Accepted November 2002

if there is still a credible basis for differentiation among Big 5 accounting firms, based on their industry clienteles.[1]

We begin the analysis by considering two contrasting ways of conceptualizing the operations of Big 5 accounting firms. The "firm-wide" perspective considers an accounting firm's practice in aggregate, typically at the country-level of analysis. The firm is the focal point and no differentiation is presumed to exist across the individual practice offices of the firm. Firm-wide studies using country-level data have documented that Big 5 audit fees are higher than the fees of non-Big 5 accounting firms (Francis 1984; Francis and Stokes 1986; Francis and Simon 1987; Chan et al. 1993; Craswell et al. 1995; DeFond et al. 2000).

An alternative, which we term the "office-level" perspective, views each individual practice office in the Big 5 network as a unique and relevant unit of analysis in its own right because audit contracting is conducted through local offices, audit engagements are administered by an audit team typically located in an office in the same city as the client's headquarters, and audit reports are issued on office-specific letterhead of the Big 5 engagement office administering the audit (Wallman 1996; Penno and Walther 1996; Francis et al. 1999; Reynolds and Francis 2000).

The implication of each perspective for audit pricing is as follows. Under the firm-wide perspective, industry expertise is a firm-wide phenomenon and therefore a firm-wide measure of industry expertise is relevant for the study of audit pricing. Under the office-level perspective, industry expertise resides in the unique human capital and experience of professionals in each office, therefore an office-level measure of industry expertise is relevant for the study of audit pricing. In both the firm-wide and office-level perspectives, the assumption is that some clients demand a higher level of industry expertise from their auditor and are willing to pay a premium for the expertise. The difference in the two perspectives is whether industry expertise is viewed as a firm-wide or office-specific phenomenon.[2]

Like other empirical studies, we assume that Big 5 industry expertise can be inferred from a firm's market share of industry audit fees. The firm-wide perspective views national market share as the relevant basis for determining Big 5 industry expertise, while the office-level perspective views city-specific (local office) industry market share as the relevant basis for determining industry expertise. The firm-wide analysis in our study uses country-level data in Australia to determine national market shares of Big 5 firms, and the office-level analysis determines industry market shares of Big 5 firms for each unique city-industry combination in the study (Adelaide, Brisbane, Melbourne, Perth, and Sydney). In both cases a larger market share implies greater industry expertise, *ceteris paribus*. The two approaches can result in quite different market share measures and rankings. For the Australian data in our study, the nationally top-ranked accounting firm in an industry is the city-specific industry leader in only 44 percent of city-industry combinations, and Francis et al. (1999) report similar results using U.S. data.

Prior studies use arbitrary market share thresholds (typically 10–20 percent) and apply these percentages across all industries to denote industry experts (e.g., Pearson and

---

[1] One explanation for result in Ferguson and Stokes (2002) is that the accounting firm mergers in the late 1980s and 1990s which reduced the Big 8 to the Big 5 have equalized market shares to the point where no meaningful differentiation exists among Big 5 auditors based on clienteles and that they are all now viewed as experts in all industries.

[2] Evidence of higher quality audits by industry specialists is supported by recent evidence that industry specialists constrain earnings management (Zhou and Elder 2002; Krishnan 2003), and result in a higher market valuation of earnings (Balsam et al. 2002). Another viewpoint is that specialization creates production economies that will reduce audit fees if efficiencies are passed on to clients (Craswell et al. 1995, 301). While plausible, there is no evidence in the research literature of lower fees for Big 5 firms that are industry specialists.

Trompeter 1994; Craswell et al. 1995; Hogan and Jeter 1999; DeFond et al. 2000; Ferguson and Stokes 2002). Our study takes a different approach and uses Big 5 industry rankings based on market shares within each industry to denote industry experts.[3] Ferguson and Stokes (2002) use this approach and find no evidence of a premium for the top-ranked auditor in each industry based on national market share data. However, we do not know *a priori* if the market perceives only the top-ranked auditor as having differential industry expertise relative to other Big 5 firms. For example, if the second-ranked firm is also perceived by the market to be an industry expert, then a model testing only the top-ranked firm would be misspecified.

Since we have no theoretical or empirical basis for knowing how many auditors are viewed as industry experts, we approach the problem as an empirical one and undertake the following series of tests. We begin by comparing the top-ranked firm in the industry with all other Big 5 auditors. The test is conducted using both firm-wide (national) data and city-specific (office) data to denote industry leaders. We then repeat the above analysis and test if each of the top two auditors in an industry (national and city) has a premium relative to other Big 5 auditors. We extend this logic and test each of the top three firms in the industry (national and city) relative to the other Big 5 auditors. We did not proceed with further testing because it was clear that, on average, only the top two firms (at most) are perceived by the market to be industry experts as evidenced by higher audit fees.

Our firm-wide analysis indicates that on average each of the top two auditors nationally has a fee premium relative to other Big 5 auditors, but the third-ranked firm does not have a premium. Our city-specific analysis indicates that, on average, only the top-ranked Big 5 auditor in an industry has a fee premium, with no premium for the second- or third-ranked firms.[4]

In an attempt to reconcile the firm-wide and office-level results, we test the joint effect of firm-wide industry expertise and office-specific industry expertise on audit fees. These results indicate that Big 5 auditors earn an average fee premium of 24 percent when they are one of the top two firms nationally *and* are also the city-specific industry leader. We further document that the top two firms nationally *do not* earn a fee premium if they are not also the city leader. Thus, the evidence shows that an auditor must be a city-level industry leader to earn a premium for industry expertise, and our study supports the office-level perspective in the pricing of Big 5 industry expertise.

The remainder of the study is organized as follows. The next section develops in more detail the firm-wide and office-level perspectives on the operations of Big 5 accounting firms and the development of industry expertise. Section III describes the data and research design. Primary results are reported in Section IV and sensitivity analyses and robustness checks are reported in Section V. Results of the study and implications for future research are discussed in the concluding section.

---

[3] Our reasoning draws on research in marketing and economics that market share is a signal of quality, and industry leadership provides the strongest signal of quality with respect to the accounting firm's industry expertise (Smallwood and Conlisk 1979; Caminal and Vives 1996; Hellofs and Jacobson 1999).

[4] At the country level, the average market shares across the 24 industries in the study are 50 percent for the top-ranked firm, 20 percent for the second-ranked firm, and 12 percent for the third-ranked firm. At the city level, the top-ranked Big 5 auditor in each city-industry combination has an average market share of 66 percent, the second-ranked firm 18 percent, and the third-ranked firm 7 percent. For both national- and city-specific measures, the differences in market shares are statistically significant between the first- and second-ranked firms, and between the second- and third-ranked firms.

II. TWO VIEWS OF BIG 5 ACCOUNTING FIRMS

The Big 5 accounting firms are organized as country-level partnerships, and have also created international organizations that bring all of the country-level partnerships together under one administrative structure. Andersen Worldwide was the prototype for these international umbrella organizations. However, while the Big 5 firms have country-level and international organizational structures, they operate through a decentralized network of semi-autonomous practice offices (Narayanan 1995; Wallman 1996; Francis et al. 1999; Chan et al. 2002). National offices are small and serve as an administrative unit for the firm. Individual engagement partners in practice offices contract for audits (in the name of the firm) with clients that are generally headquartered in the same locale, and administer the audit engagement even though other offices may be involved. The final audit report is signed by the engagement partner on the letterhead of the local office of the Big 5 accounting firm administering the engagement.[5] Descriptively, then, audit contracting and audit reporting occurs at the office-level of each Big 5 accounting firm even though the firms have national and international organizational structures.

In predicting audit pricing the central question is the degree to which positive network externalities exist across individual offices of Big 5 accounting firms.[6] Positive network externalities will exist if an accounting firm develops a firm-wide reputation and if individual offices within a Big 5 firm's network capture the benefit of the firm-wide reputation. Craswell et al. (1995) argue that Big 5 firms have developed brand name reputations. Positive network externalities are likely for brand name reputations, and all offices of a Big 5 firm should benefit from the firm's reputation. Some of the factors that contribute to this are firm-wide quality control procedures (e.g., standardized audit programs, peer reviews) and uniform firm-wide personnel practices with respect to hiring, training, and promotions. The finding of a Big 5 premium in many countries is consistent with positive network externalities across individual offices for a Big 5 accounting firm's brand name reputation.

With industry expertise, it is less clear if there are positive firm-wide externalities. We now develop in greater detail the underlying assumptions about industry expertise based on the firm-wide and office-level views of the operations of Big 5 accounting firms.

Reasons Why Industry Expertise Is Firm-Wide

The firm-wide view assumes that industry expertise is captured by the firm as whole and therefore exists at the firm level. Expertise in individual offices is assumed to be uniform across offices and the national office facilitates the transfer of expertise across offices so that clients can have seamless access to the firm-wide expertise they demand. In this view, all offices of a particular firm are suppliers of homogenous expertise. Uniform industry expertise is provided across individual offices through standardized national training programs, standardized audit programs, travel by industry experts between offices and, with advances in technology through knowledge management systems, video-conferencing and other contact between local office partners and industry experts in other offices.

This view is supported by the fact that the Big 5 firms hold themselves out through the media and their web sites as having firm-wide operations organized along industry lines, standardized national training, and standardized audit programs. If the firm-wide view is descriptive of how accounting firms operate—one in which there are strong positive network externalities—then national (firm-wide) industry expertise is all that will matter in the pricing of audits, and clients that demand a higher level of industry expertise will pay more to the Big 5 firms that are perceived to have such expertise. Under this view we assume clients use an auditor's industry ranking based on national market share as the basis for determining auditor industry expertise.

---

[5] Francis et al. (1999) report that 97 percent of U.S. audit reports of Big 5 firms are issued by offices located in the same geographic locale as the client's headquarters. In our Australian sample, all companies are headquartered in the same city as the local office of the Big 5 accounting firm administering the audit and issuing the audit report.

[6] See Katz and Shapiro (1985) and Bental and Spiegal (1995) for a general discussion and economic analysis of network externalities.

Reasons Why Industry Expertise Is Office-Specific

Accounting is a knowledge-based professional service, and under the office-level perspective industry expertise is seen to reside in the unique personnel of the firm and is acquired through experience working with individual clients in specific industries (Solomon et al. 1999). This argument is also consistent with Carcello et al. (1992) and Behn et al. (1997) who find that controllers of *Fortune 1000* companies value highly the industry experience of the primary audit engagement team. Under the office-level perspective it is unlikely there are strong positive network externalities related to industry expertise. The reason is that the accounting firm's investment in industry expertise is really an investment in the human capital of its professional staff, and the expertise derived from these human capital investments tends to be both client- and office-specific and, therefore, is not readily captured by the firm as a whole (which the firm-wide view assumes is possible). Gilson and Mnookin (1985) make similar arguments with respect to human capital in large law firms. Some aspects of industry expertise can be captured by Big 5 firms and distributed to local offices, for example, industry-tailored audit programs and databases with best industry practices. But deep industry knowledge resides in the firm's individual experts and is limited to the specific engagements on which these individuals work, and most auditors continue to service clients predominantly in one locale. Under this view we assume clients use an auditor's industry ranking based on their office-level (city-specific) market share as the basis for determining auditor industry expertise.

In sum, a local office of a Big 5 accounting firm will have a more credible claim of industry expertise to their clients (or prospective clients) if they are a leading firm in that industry for that particular city. The fact that an accounting firm has a "large" national market share counts for relatively little in city-specific audit markets if the national market share statistics are driven by clients located primarily in other cities. It turns out that this is the typical case. Francis et al. (1999) document that the national industry leader is the city-specific industry leader in only 25 percent of cities in the United States; in our Australian sample, the national industry leader is the city-specific industry leader in only 44 percent of city-industry combinations. If the local-office view is correct—one in which positive network externalities do not exist—then local-office industry expertise is what will be priced in audits for those clients demanding industry expertise.[7]

---

[7] Empirically we know that industries are not randomly distributed. For U.S. data from 1990, Francis et al. (1999) find that the four-city concentration ratio is 37 percent for 27 industry sectors. This means that 37 percent of companies in an industry are located in the four cities with the most companies. Similarly, for Big 5 industry leaders, over half of their industry clientele is located in the four offices of the firm having the most clients in the industry. Thus, industries tend to cluster in certain cities, and auditor market shares also cluster in these same cities, which suggest that a small set of offices can drive the firm's national market share. However, the question remains whether a firm can achieve positive firm-wide externalities for industry expertise across all of its offices.

III. DATA AND MODEL SPECIFICATION

Sample

We start with the sample of 1,084 Australian publicly listed companies in Ferguson and Stokes (2002). Their sample represents nearly all Australian publicly listed companies for fiscal 1998 and is collected from the following sources: Huntleys Datanalysis, Huntleys Datanalysis (Delisted Companies), the Australian Stock Exchange Library, the Australian Stock Exchange Microfiche Collection in the State Library of New South Wales, Connect 4, and Reuters Beacon.

Audit fees are examined for companies headquartered in the five largest Australian cities: Adelaide, Brisbane, Melbourne, Perth, and Sydney. These five cities contain the headquarters of 1,046 out of 1,084 companies in the Ferguson and Stokes (2002) sample, and are audited by offices of accounting firms located in the same city (based on information in the audit report). The other 38 publicly listed companies are located in several smaller regional cities, are audited by offices in these locales, and are thus excluded from the study.

The sample of 1,046 companies is distributed as follows: Adelaide has 43 observations, Brisbane has 93 observations, Melbourne has 223 observations, Perth has 290 observations, and Sydney has 397 observations. As in Craswell et al. (1995) and Ferguson and Stokes (2002), to avoid the confounding effect of Big 5 brand name on audit fees, the regression models are estimated for the 681 companies in the sample having Big 5 auditors, though industry market share calculations and rankings are based on the total sample of 1,046 companies.[8]

Audit fees measure industry output and can be used to calculate Big 5 market shares in Australia because fees are publicly disclosed in corporate annual reports.[9] For convenience, the names of the Big 5 accounting firms are abbreviated as follows: Arthur Andersen (AA), Deloitte Touche (DT), Ernst & Young (EY), KPMG (KPMG), and PricewaterhouseCoopers (PWC). Industry rankings are based on an accounting firm's share of total audit fees, measured nationally for the firm-wide perspective and measured for each industry per specific city for the office-level perspective. The industry groupings used in Craswell et al. (1995) and Ferguson and Stokes (2002) are also used in this study and are based on the 24 categories used by the Australian Stock Exchange.

Table 1 provides market shares of each Big 5 firm. Panel A of Table 1 shows the total national market share of each Big 5 firm and the total city-level market share of each Big 5 firm. Nationally, the two dominant firms are KPMG with 28 percent of total audit fees followed closely by PWC with 24 percent, a difference of only 4 percent. Panel A shows that national market shares are more evenly distributed among the Big 5 firms than city market shares. National data show a relatively narrow range from 28 percent for the top-ranked firm to 10 percent for the fifth-ranked Big 5 firm, a difference of only 18 percent. By contrast, three different firms are city leaders and fees average 38 percent for the city leader compared to 26 percent for the second-ranked firm. In addition, the range from top

---

[8] National industry rankings of auditors are unaffected if using the full sample of 1,084 observations (which includes the 38 observations in smaller regional cities and represents less than 1 percent of total audit fees for the full sample of 1,084 observations).

[9] As in prior studies, all audit fees are attributed to the city-specific office that contracts for the audit and issues the audit report signed on the local office letterhead of the Big 5 firm (Francis et al. 1999; Reynolds and Francis 2000). We recognize this introduces two potential sources of error in measuring city-level audit market shares. First, it ignores fees from private companies, though these fees are unlikely to be large relative to publicly listed companies. Second, it potentially overstates local office fees to the extent other offices participate in engagements, although, the primary reputation effect is most likely to accrue to the engagement office that administers the audit and issues the final audit report. Therefore, while we believe it is reasonable to attribute all fees to the office administering the audit, we recognize the potential for measurement error.

TABLE 1
City and National Market Shares Based on 1998 Australian Audit Fees
(percentages of audit fees in parenthesis)

Panel A: Overall City-Level and National-Level Market Shares

Perth n = 290	Sydney n = 397	Melbourne n = 223	Brisbane n = 93	Adelaide n = 43	National n = 1084
EY (33)	EY (25)	KP (41)	PW (45)	KP (45)	KP (28)
PW (21)	PW (21)	PW (27)	KP (26)	DT (33)	PW (24)
KP (17)	KP (18)	AA (13)	AA (08)	PW (09)	EY (16)
AA (11)	AA (14)	DT (09)	EY (08)	EY (04)	AA (12)
DT (04)	DT (11)	EY (04)	DT (04)	AA (01)	DT (10)
NB5 (14)	NB5 (11)	NB5 (06)	NB5 (09)	NB5 (08)	NB5 (10)

Panel B: City-Level and National-Level Industry Leaders

Australian Stock Exchange Industry	Perth	Sydney	Melbourne	Brisbane	Adelaide	National #1	National #2
01 Mining–Gold	EY (33)	PW (47)	PK (34)	EY (27)	DT (99)	PW (21)	DT (20)
02 Mining–Other Metals	KP (43)	PW (40)	EY (52)	PW (96)	DT (100)	PW (57)	KP (19)
03 Solid Fuels	—	—	AA (47)	—	—	AA (47)	PW (45)
04 Energy	EY (39)	KP (36)	PW (58)	PW (47)	KP (100)	KP (33)	PW (18)
05 Infrastructure and Utilities	—	DT (73)	AA (75)	AA (42)	PW (100)	DT (43)	AA (28)
06 Developers and contractors	KP (62)	KP (48)	DT (53)	KP (35)	KP (50)	KP (45)	PW (13)
07 Building Materials	PW (44)	AA (33)	PI (100)	EY (100)	KP (100)	AA (29)	DT (26)
08 Alcohol and Tobacco	BD (100)	PW (73)	PW (98)	PW (81)	PW (67)	PW (86)	BD (03)
09 Food and Household	FJ (63)	EY (54)	PW (30)	KP (71)	—	EY (50)	KP (31)
10 Chemicals	—	PK (52)	KP (97)	KP (100)	HC (100)	KP (77)	PK (11)

(continued on next page)

TABLE 1 (continued)

Australian Stock Exchange Industry	Perth	Sydney	Melbourne	Brisbane	Adelaide	National #1	National #2
11 Engineering	EY (53)	EY (66)	PK (51)	PW (69)	EM (44)	EY (47)	PW (16)
12 Paper and Packaging	—	BD (100)	KP (83)	—	—	KP (81)	PW (14)
13 Retail	EY (94)	EY (31)	PW (74)	AA (54)	PW (71)	PW (44)	EY (23)
14 Transport	—	DT (72)	KP (79)	—	KP (100)	DT (66)	KP (27)
15 Media	PW (100)	AA (64)	AA (88)	—	PK (100)	AA (63)	EY (12)
16 Banks	KP (85)	PW (50)	KP (100)	KP (97)	EY (100)	KP (53)	PW (22)
17 Insurance	KP (100)	EY (52)	KP (49)	—	—	EY (36)	PW (22)
18 Telecommunications	PW (84)	DT (36)	PW (95)	PW (91)	—	PW (80)	DT (07)
19 Investment and Financial Services	PW (52)	KP (28)	DT (28)	DT (37)	DT (71)	KP (23)	DT (17)
20 Property Trusts	EY (100)	EY (40)	MS (23)	KP (100)	BD (77)	EY (37)	PW (27)
21 Healthcare and Biotechnology	KP (70)	PW (46)	KP (73)	JR (74)	DT (100)	KP (48)	DT (17)
22 Miscellaneous Industrials	KP (28)	PW (24)	DT (56)	KP (33)	DT (46)	DT (26)	PW (19)
23 Diversified Industrials	EY (48)	KP (44)	KP (87)	EY (78)	KP (100)	KP (70)	PW (12)
24 Tourism and Leisure	PW (60)	PW (35)	EY (84)	AA (40)	—	EY (35)	PW (25)

AA—Arthur Andersen
BD—BD Nelson Parkhill
DT—Deloitte Touche
EM—Edwards Marshall
EY—Ernst & Young
FJ—FA Jones
HC—Hall Chadwick
KP—KPMG
JR—Johnston Rorke
MS—Moore Stephens
NB5—a non-Big 5 Auditor
PI—Pitcher Partners
PK—Pannell Kerr Foster
PW—PricewaterhouseCoopers

to bottom at the city level averages 33 percent (compared to 18 percent at the national level).

Panel B of Table 1 reports industry market shares by the top two firms nationally, and the city-specific leaders for each of the 24 industry categories of the Australian Stock Exchange. Even though KPMG and PWC are the two leading firms in overall market share and earn 52 percent of total fees in the sample, national-level industry leadership is quite diverse. KPMG or PWC are the top-ranked firm in only 13 of 24 industries, and are the second-ranked firm in 15 of 24 industries. Thus, another firm is one of the top two firms in an industry in 20 of 48 instances.

There are a potential of 120 city-specific industry leaders in Panel B (24 industries times 5 cities). However, for 17 city-industry combinations there are no observations and therefore no industry leaders. For these 103 city-industry combinations in the study the Big 5 national leader is the city-level industry leader in only 45 cases (44 percent). In addition, for 13 of the city-industry combinations a non-Big 5 accounting firm is the industry leader.[10] This leaves 90 city-industry combinations for which a Big 5 firm is the industry leader and these are distributed as follows: AA has 8 industries in which it is the city-specific market leader, DT has 12 industries, EY has 17 industries, KPMG has 27 industries, and PWC has 26 industries. Again, even though KPMG and PWC are the two dominate firms nationally, they are city-level leaders in only 53 of the 90 city-industry combinations having a Big 5 firm as city leader.

In sum, the national industry leader is the city-level industry leader in less than half of the industry-city combinations in Panel B (45 of 103 cases). No industries have the same accounting firm as the industry leader for all five cities, and 16 of the 24 industries have three (or more) different accounting firms as market leaders for the five cities. Thus, there is no general pattern in which individual accounting firms dominate industries across all cities.

Model Specification

A cross-sectional audit fee regression model similar to that used in prior audit fee research is used to estimate the effects of industry expertise on audit fees (Francis 1984; Francis and Stokes 1986; Craswell et al. 1995; Ferguson and Stokes 2002). Audit fee regression models use a set of variables to control for cross-sectional differences in factors that affect fees such as client size, audit complexity, and auditor-client risk sharing. These models have good explanatory power (adjusted R^2s of 0.70 and higher) and have been robust across different samples, time periods, countries, and sensitivity analyses for model misspecification (Francis and Simon 1987; Chan et al. 1993).

Two auditor variables are specified to test the effect of industry expertise on audit pricing: the first variable, *AUDITOR#1*, indicates if the observation is audited by the specific Big 5 firm that is industry leader for that particular observation's industry. The second variable, *AUDITOR#2*, indicates if the observation is audited by second-ranked Big 5 firm in that particular industry. The default comparison group consists of observations audited by all other Big 5 auditors. This is similar to the research design in prior audit fee research in which an indicator variable is used to test if Big 5 auditors earn a fee premium over non-Big 5 auditors. The design tests for differential audit fees through a cross-sectional comparison of audit fees of the top two industry leaders with other firms, after controlling for other factors affecting fees. Formally, the test determines if there is a significant positive intercept shift (higher fees) in the fitted regression model for observations audited by either of the two top-ranked accounting firms in the industry.[11]

---

[10] These 13 city-industry combinations with a non-Big 5 market leader represent only 27 of the 1,046 companies in the study, and only 0.8 percent of total audit fees for the full sample of 1,046 companies.

An OLS regression model is estimated pooling observations with Big 5 auditors and company headquarters in Adelaide, Brisbane, Melbourne, Perth, and Sydney. Observations are pooled across cities to increase statistical power, though a sensitivity analysis of individual cities is reported later. The model is specified in two ways: the first estimation defines the top two accounting firms based on national-level market share data, and the second estimation defines the top two accounting firms using city-level market share data, which means that an industry can have different auditors as industry experts in different cities depending on auditor industry rankings in each city.

The OLS regression model is specified as follows:

$$\begin{aligned} LAF = {} & b_0 + b_1 LTA + b_2 LSUB + b_3 CATA + b_4 QUICK + b_5 DE + b_6 ROI \\ & + b_7 FOREIGN + b_8 OPINION + b_9 YE + b_{10} LOSS + b_{11} AUDITOR\#1 \\ & + b_{12} AUDITOR\#2 + e \end{aligned}$$

where:

LAF = natural log of audit fees;
LTA = natural log of total assets;
$LSUB$ = natural log of the number of subsidiaries;
$CATA$ = ratio of current assets to total assets;
$QUICK$ = ratio of current assets (less inventories) to current liabilities;
DE = ratio of long-term debt to total assets;
ROI = ratio of earnings before interest and tax to total assets;
$FOREIGN$ = proportion of subsidiaries that represent foreign operations;
$OPINION$ = indicator variable, 1 for qualified audit report;
YE = indicator variable, 1 for non-June 30th year-end;
$LOSS$ = indicator variable, 1 for loss in any of the past three years;
$AUDITOR\#1$ = indicator variable, 1 for top-ranked auditor in the industry;
$AUDITOR\#2$ = indicator variable, 1 for second-ranked auditor in the industry; and
e = error term assumed to have normal OLS regression properties.

As in prior studies transformations are made to linearize certain variables (*LAF*, *LTA*, *LSUB*). Observations having a zero value for *LSUB* are re-coded to a small positive value (.00001) to enable a logarithmic transformation. Three additional variables, *QUICK*, *DE*, and *ROI* are windsorized to a maximum value of ± 3 standard deviations to control for extreme values.

With respect to the ten control variables in the model, higher fees are expected (positive signs) for larger clients (*LTA*), for greater audit complexity (*LSUBS* and *FOREIGN*), and for greater audit risk (*DE* and *CATA*). A positive sign is also expected for *OPINION* as prior studies document higher fees associated with modified opinions, possibly due to more investigative efforts in such circumstances. Negative signs are expected for the variables *QUICK*, *ROI*, *LOSS*, and *YE*. *QUICK* is an audit risk variable and clients with a smaller quick ratio are riskier (less liquid) and therefore are expected to have higher audit fees. Prior studies find that clients with higher *ROI* have lower fees, which is consistent with auditor-client risk sharing, i.e., more profitable clients pose less risk to the auditor resulting in lower fees. However, prior studies also show that clients with operating losses (*LOSS*) have lower fees, possibly due to an inability to pay higher fees. Finally, clients with off peak non-June 30th fiscal year-ends (*YE*) are expected to have lower fees.

---

[11] As explained in the introduction, we undertook a series of tests at the national- and city-level of analysis to determine if the top-ranked firm has a premium, if each of the two top-ranked firms has a premium, and if each of the three top-ranked firms has a premium, relative to other Big 5 auditors. On the basis of these analyses it appears that only the top two ranked firms (nationally) have a premium, and the models reported in our study are based on this analysis. The models testing only the top-ranked firm, and the models testing each of the three top-ranked firms, are available upon request.

Descriptive Statistics

Descriptive statistics are reported in Table 2 for the full sample of 1,046 observations, and for the 681 observations audited by Big 5 accounting firms in 1998. The 681 observations with Big 5 auditors represent 65 percent of the 1,046 observations and 91 percent of total audit fees. At 65 percent, the Big 5 share of clients in Australia is somewhat smaller than the U.S., and this may be due to the smaller average size of companies in Australia.

TABLE 2
1998 Australian Sample Descriptive Statistics
(June 30, 1998 exchange rate $1 AUD = .6135 USD)

	Full Sample (n = 1,046)		**Big 5 Sample (n = 681)**	
Variables	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
LAF	3.82	1.43	4.16	1.47
LTA	10.56	2.30	11.08	2.39
LSUB	−0.66	5.39	−0.37	5.43
CATA	0.36	0.27	0.35	0.27
Quick	3.75	5.83	3.20	5.30
DE	0.11	0.15	0.12	0.15
ROI	−0.10	0.32	−0.08	0.31
Foreign	0.17	0.27	0.18	0.26
Opinion	13%		11%	
YE	20%		23%	
Loss	61%		56%	
Big 5	65%			
City IL			36%	

LAF = natural log of total audit fees;
LTA = natural log of total assets ($000);
LSUB = natural log of the number of subsidiaries;
CATA = ratio of current assets to total assets;
Quick = ratio of current assets (less inventories) to current liabilities;
DE = ratio of long-term debt to total assets;
ROI = ratio of earnings before interest and tax to total assets;
Foreign = proportion of subsidiaries that are foreign;
Opinion = 1 if qualified audit opinion;
YE = 1 if non-30th June balance date;
Loss = 1 if reported loss in any of the last three years;
Big 5 = 1 if Big 5 auditor; and
City IL = 1 if auditor is the Big 5 City Industry Leader.

However, the Big 5 firms do earn over 90 percent of total audit fees. The Big 5 percentages of clients (fees) for the five cities are as follows: Adelaide, 70 percent (92 percent); Brisbane, 69 percent (91 percent); Melbourne, 62 percent (94 percent); Perth, 69 percent (86 percent); and Sydney, 66 percent (89 percent).

IV. RESULTS

The study's primary results are presented in Table 3. Model (1) is the estimation using national-level data to determine the top two accounting firms in each industry, and Model (2) is the estimation using city-level data to determine the top two accounting firms in each unique city-industry combination. Both models are significant at $p < .01$ and have adjusted R^2s of 0.80, which are consistent with prior research and indicate good explanatory power for the models. The control variables are significant in the expected direction with the exception of *OPINION*, *LOSS*, and *YE*, which are not significant in any of the models.

Model (1) shows that both the national industry leader and the second-ranked firm in the industry have a premium over other auditors. The coefficient on the industry leader is significant at $p = .069$, and the second-ranked firm is significant at $p = .03$. An F-test cannot reject the null hypothesis of no difference in these two parameter estimates. We conclude from Model (1) that both of the top two nationally ranked firms in an industry ($n = 351$) have a comparable premium over the audit fees of other Big 5 auditors ($n = 330$). This result differs from Ferguson and Stokes (2002) who found no fee premia based on national market shares. However, they used a different model specification and conducted tests using market share threshold percentages to define industry experts (10, 15, and 20 percent) or tested the top-ranked firm alone.

Model (2) shows that the city-level industry leader has a fee premium over other auditors, but the second-ranked firm does not have a premium. The coefficient on the city-level industry leader is significant at $p = .01$. We conclude from Model (2) that the city-level industry leader ($n = 247$) has a fee premium relative to other Big 5 auditors ($n = 434$).

The results in Model (1) provide support for the firm-wide view of industry expertise, while the results in Model (2) provide support for the office-level view. We now attempt to reconcile these results by examining the potential joint effects of firm-wide and office-level industry expertise on the pricing of audits. Model (3) in Table 3 creates three test variables: (1) an indicator variable if the auditor is one of the two top-ranked accounting firms in an industry nationally, and is also the city-specific industry leader ($n = 169$); (2) an indicator variable if the auditor is one of the two top-ranked firms in an industry nationally, but is *not* the city-specific industry leader ($n = 182$); and (3) an indicator variable if the auditor is the city-specific industry leader, but is *not* one of the two top-ranked firms in an industry nationally ($n = 78$). The default comparison group is companies whose Big 5 auditor is neither the city-level industry leader nor one of the two top-ranked firms nationally ($n = 252$).

Model (3) shows that auditors have significantly higher fees when they are both the city-specific leader and one of the top two firms nationally ($p = .001$). The coefficient on the auditor variable is 0.213, and using the procedure described in Craswell et al. (1995, 307), this parameter value represents an average audit fee premium of 24 percent.[12] The other two auditor indicator variables in Model (3) are not significant at conventional levels.

---

[12] The procedure calculates the percentage effect of the intercept shift on the dependent variable (natural log of audit fees), and is defined as $e^z - 1$, where z is the parameter value for auditors who are market leaders.

TABLE 3
Audit Fee Regression Models
(dependent variable is natural log of audit fees)

		National #1 and #2 Firms			City #1 and #2 Firms			Combined Model			Model 3 excluding Adelaide, Brisbane		
		Model 1			Model 2			Model 3			Model 4		
Control Variables	**Exp. Sign**	**Estimate**	**t-value**	**Prob***	**Estimate**	**t-value**	**Prob***	**Estimate**	**t-value**	**Prob***	**Estimate**	**t-value**	**Prob***
Intercept		−1.54	−8.05	0.001	−1.52	−7.99	0.001	−1.45	−7.63	0.001	−1.43	−6.97	0.001
LTA	+	0.47	31.20	0.001	0.46	30.90	0.001	0.46	30.84	0.001	0.46	28.70	0.001
LSUB	+	0.05	9.20	0.001	0.05	9.04	0.001	0.05	9.10	0.001	0.05	8.25	0.001
CATA	+	0.94	8.77	0.001	0.94	8.82	0.001	0.93	8.81	0.001	0.97	8.44	0.001
Quick	−	−0.03	−5.06	0.001	−0.03	−5.14	0.001	−0.03	−5.12	0.001	−0.03	−5.11	0.001
DE	+	0.76	4.21	0.001	0.75	4.16	0.001	0.79	4.38	0.001	0.77	3.96	0.001
ROI	−	−0.47	−4.50	0.001	−0.45	−4.37	0.001	−0.46	−4.45	0.001	−0.45	−4.37	0.001
Foreign	+	0.69	6.60	0.001	0.70	6.71	0.001	0.69	6.60	0.001	0.73	6.43	0.001
Opinion	+	0.01	0.04	0.968	0.01	0.07	0.946	0.01	0.15	0.885	−0.01	−0.03	0.973
YE	−	0.03	0.44	0.663	0.03	0.42	0.678	0.01	0.18	0.857	0.04	0.54	0.586
Loss	−	−0.01	−0.15	0.881	−0.01	−0.10	0.919	0.01	0.02	0.984	−0.03	−0.36	0.719

(continued on next page)

TABLE 3 (continued)

		National #1 and #2 Firms			City #1 and #2 Firms			Combined Model			Model 3 excluding Adelaide, Brisbane		
		Model 1			Model 2			Model 3			Model 4		
Experimental Variables	**Exp. Sign**	**Estimate**	**t-value**	**Prob***	**Estimate**	**t-value**	**Prob***	**Estimate**	**t-value**	**Prob***	**Estimate**	**t-value**	**Prob***
Industry Leader		0.094	1.48	0.069	0.143	2.34	0.010						
Second Largest Firm		0.117	1.88	0.030	0.050	0.77	0.222						
National #1 or #2, and city #1								0.213	3.16	p = .001	0.240	3.18	p = .001
National #1 or #2, not city #1								−0.018	−0.28	p = .392	−0.029	−0.43	p = .332
City #1, not National #1 or #2								−0.081	−0.96	p = .169	−0.044	−0.46	p = .323
			Model 1			**Model 2**			**Model 3**			**Model 4**	
F-statistic (p-value)			227 (<.001)			227 (<.001)			213 (<.001)			189 (<.001)	
Adjusted R^2			0.80			0.80			0.80			0.81	
Sample Size			n = 681			n = 681			n = 681			n = 587	

---

* All p-values are two-tailed, except the experimental auditor market share variables, which are one-tailed.
LTA = natural log of total assets;
LSUB = natural log in the number of subsidiaries;
CATA = ratio of current assets to total assets;
Quick = ratio of current assets (less inventories) to current liabilities;
DE = ratio of long-term debt to total assets;
ROI = ratio of earnings before interest and tax to total assets;
Foreign = proportion of subsidiaries that represent foreign operation;
Opinion = indicator variable, 1 = qualified audit report;
YE = indicator variable, 1 = non-June 30th year end;
Loss = indicator variable, 1 = loss in any of the past three years.

Model (3) is re-estimated dropping those observations in Adelaide and Brisbane, the two smallest cities in the sample. This result is reported as Model (4) in Table 3 and is qualitatively the same as Model (3), indicating that the results in Model (3) are not driven by potentially smaller industry sizes and auditor clienteles in Adelaide or Brisbane. The auditor coefficient is slightly larger (.24) equating to an average premium of 27 percent for auditors that are both city leaders and one of the top two firms nationally. As a further sensitivity analysis, Model (3) is also re-estimated dropping each of the five cities one at a time. Despite the loss of observations, these additional results are qualitatively the same as Model (3) and auditors that are both the local industry leader and one of the top two firms nationally have a significant fee premium ($p < .05$) in all cases, and the other two auditor indicator variables are insignificant. To be sure that individual industries do not drive the results, Model (3) is re-estimated deleting each of the 24 industries one at a time. The fee premium for auditors that are one of the two top-ranked firms nationally and also the city-specific industry leader is robust to dropping individual industries, and the one-tail p-values range from .001 to .034 and average .005.

At face value the results in Models (3) and (4) imply that both firm-wide and office-level expertise affects audit pricing and that neither view alone is correct or sufficient to explain a fee premium for industry expertise. However, further analysis shows that in fact the results are driven by office-level industry leadership. We begin by comparing the city-level market shares of the two top-ranked firms nationally when they are also the city industry leader ($n = 169$) vs. when they are not the city industry leader ($n = 182$). The purpose is to see if relative market share data help to explain why a premium is earned by national leaders only when they are also city leaders. When the two top-ranked firms nationally are also the city-industry leader, they have an average market share of 66 percent, which is significantly greater than the average market share of 17 percent for the second-ranked firm in the city (t-statistic 32.5, $p < .001$). However, when the two top-ranked firms nationally are *not* the city-industry leader, their average market share is only 17 percent, which is significantly less than the average market share of 48 percent for the city-industry leader in those cities (t-statistic -30.1, $p < .001$). If expertise is inferred by having a leading local-office market share, then the two top-ranked firms nationally clearly have significantly smaller market shares than the city leader in those cities where they are not the city leader. These market share comparisons help to better understand why an audit premium is earned only when the firm is also the city-level industry leader.[13]

Next we show that national industry leadership itself is a statistical artifact of the city-specific offices where the firm is the industry leader. The nationally top-ranked Big 5 firm in each industry is the city leader in 45 of 103 city-industry combinations in the study (recall from Table 1, Panel B there are 120 possible city-industry combinations in the study, less 17 that have no observations). In aggregate, the national Big 5 industry leaders' audit fees for these 45 city-industry combinations in which the national leader is also the city leader represent 95 percent of the national leaders' total industry audit fees. Importantly, the same Big 5 firm would still be the nationally top-ranked firm in 21 of 24 industries in the study if the top-ranked accounting firm's fees were based solely on those offices where it is also the city-specific industry leader. This analysis illustrates why the national market share data for the top-ranked firm is a statistical artifact of the city-specific practice offices in which these firms are industry leaders. In other words, national market share data for the top firm nationally is driven almost exclusively by the individual offices where the firm is also the city leader.[14]

---

[13] The results are comparable when dropping those city-industry combinations in which the city leader has only one or two clients. This is important because it rules out that the market share differences are not driven by small industries or small auditor clienteles.

In sum the top two firms nationally earn a premium *only* in those specific offices where they are also the city-industry leader, and it turns out that national-industry leadership data are really driven by the individual offices where the firm is the city leader. We conclude that the evidence does not support the existence of positive firm-wide externalities for industry expertise beyond those offices where the firm is the city-specific industry leader, and that the pricing of expertise appears to be limited to those individual offices where the auditor is the city leader.[15]

There is one anomaly in our results, however. Model (3) in Table 3 show that the city-level industry leader does *not* earn a premium if they are not also one of the top two firms nationally. This is the case for 78 observations representing 24 of the 90 city-industry combinations having a Big 5 firm as city-level industry leader. An additional analysis of these 78 observations is now undertaken to better understand why there is no fee premium. One explanation is that despite being the top-ranked firm in the city, these auditors do not have a significantly different market share from the second-ranked firm. These 78 city-industry leaders have an average market share of 63 percent, which is significantly greater than the average market share of 18 percent for the second-ranked firm in the city (t-statistic 5.19, $p < .001$). However, in eight of the 24 city-industry combinations (representing ten of 78 observations) the city leaders have a 100 percent market share. This potentially exaggerates the market share differences as there are only ten observations in these eight city-industry combinations.[16] For the remaining 16 city-industry combinations (representing 68 of the 78 observations) the city leaders' average market share is 44 percent, compared to an average market share of 28 percent for the second-ranked firm. In other words, for 68 of the 78 observations, the city-leader's market share is much closer to the second-ranked firm in the city, with an average difference of 16 percent (vs. a much larger difference of 39 percent in the 66 city-industry combinations where the city leader earns a premium). So it is possible that despite being the nominal city-level industry leader, in these 16 city-industry combinations the city leader does not stand out as being qualitatively different from the second-ranked firm, in which case there would be no perceived difference in expertise and no basis for a differential audit fee.

Finally, in terms of overall materiality, these 24 city-industry combinations in which the city leader does not earn a fee premium represent only 10 percent of total audit fees in the sample, and 29 percent of companies. By contrast the 66 city-industry combinations in which the Big 5 city leader earns a fee premium represent 90 percent of total audit fees in the sample, and 71 percent of companies. Median total assets are $87.9 million (Australian dollars) for observations in the 66 city-industry combinations, compared to only

---

[14] A similar though less dramatic effect occurs for the second-ranked firm nationally. The second-ranked Big 5 firm in each industry is the Big 5 city leader in 21 of 103 possible city-industry combinations, and representing 14 of the 24 industry categories in the study. These second-ranked firms would still be the nationally second-ranked firm in eight of the 14 industries where they are a city leader if the firms' industry fees were based solely on these offices.

[15] As a further sensitivity analysis, we test if auditors that are industry leaders in two or more cities earn higher fees than auditors that are the city-leader in only one city. The tests indicate that audit firms do not earn a relatively higher fee premium if they are the industry leader in multiple cities compared to audit firms with only one city, which means there is no evidence of positive externalities arising from multiple city leadership. This finding reinforces the importance of individual office-level industry leadership in the perception and pricing of Big 5 industry expertise.

[16] In addition, in seven of these eight city-industry combinations, the city leader has only one audit client, and this may limit the firm's ability to credibly claim expertise due to lack of multiple clients.

$22.0 million for observations in the 24 city-industry combinations. Thus, both clients and fees are much smaller in the 24 city-industry combinations where the city-leader does not earn a fee premium. This is consistent with prior evidence that an industry premium is more likely for larger companies (Craswell et al. 1995; Ferguson and Stokes 2002).[17] It is possible these smaller firms are either unable or unwilling to pay a premium to the city-level leader for its industry expertise.

In conclusion, the empirical results in Table 3 provide evidence of a Big 5 audit fee premium for industry expertise, and this result is explained primarily by office-level industry leadership in city-specific locales that represent 90 percent of total audit fees in the sample. Further implications are discussed in the concluding section but it appears that the traditional firm-wide view is not adequate to fully describe the operations of Big 5 accounting firms, at least in terms of explaining the audit market's perception and pricing of auditor industry expertise.

V. SENSITIVITY ANALYSIS AND ROBUSTNESS CHECKS

Additional sensitivity analyses are reported in this section. All tests are based on Model (3) in Table 3. The analyses demonstrate the robustness of Model (3) to a wide range of sample selection and model specification issues.

Individual Big 5 Firms

Model (3) is re-estimated dropping each Big 5 firm one at a time to be sure individual Big 5 firms do not drive the results. Despite considerably reduced sample sizes in each of these five estimations, the results are qualitatively the same as Model (3). The auditor indicator variable for observations in which the auditor is one of the national top two firms and also the city leader is significant in all cases at $p < .07$, with no evidence of a significant fee premium for the other two auditor variables.

Industry Sensitivity

An analysis is undertaken to assure that small industries do not drive the statistical results. Industry expertise is not necessarily a meaningful concept if there are only a few firms in an industry in a particular city. Model (3) in Table 3 is re-estimated dropping all industries in each city that have fewer than five observations. For this reduced sample ($n = 584$) the results are comparable to Model (3) for the full sample.

Table 1, Panel B shows that city-level industry leaders have market shares of 90 percent or more in 29 of the industry-city combinations in the study. If the premium for industry expertise is driven by this set of observations, then it could be argued that the audit fee premium represents an economic rent arising from the auditor's monopoly power rather than a premium for industry expertise. We re-estimate Model (3) in Table 3, dropping observations in these 29 industry-city combinations, and the result for this reduced sample ($n = 644$) is consistent with Model (3) for the full sample. Therefore, the audit fee premium is not driven by the auditor's potential monopoly power. We also dropped 42 observations that are city-industry leaders with only one client, to further rule out that small industries

---

[17] Other than being smaller in size, the observations in these 24 city-industry combinations are not unusual in terms of industry or auditor relative to the 66 city-industry combinations where the city leader earns a fee premium. The distributions across broad industry sectors (mining, manufacturing/distribution, services) are comparable, and the two dominant auditors are KPMG and PricewaterhouseCoopers in both groups. The 78 observations are more heavily distributed in Perth for the 24 city-industry combinations (59 percent compared to 11 percent), but we do not know why this might affect the fee premium, particularly since the industry and auditor distributions are comparable.

/clienteles do not drive the statistical results. Model (3) is unchanged when estimated without these 42 observations.

Finally, Model (3) is re-estimated dropping all observations in the mining sector (industries 01–04) to assure that this large sector of the Australian economy does not drive the statistical results. For this reduced sample (n = 456) the results are comparable to Model (3) for the full sample.

Effect of Client Size on Industry Premium

Client size is controlled by including total assets (*LTA*) in the fee model. However, it is still possible that the magnitude of the fee premium in Model (3) is sensitive to size. For example, larger clients might have greater bargaining power with auditors and therefore may pay a lower premium for industry expertise. To determine if the premium in Model (3) in Table 3 is sensitive to client size, an additional variable is added for the interaction of company size (*LTA*) with the auditor indicator variable for those observations in which the auditor is one of the top two firms nationally and also the city leader. The interaction term is positive and significant ($p < .001$). Therefore the premium in Model (3) does increase as client size increases, which is consistent with both Craswell et al. (1995) and Ferguson and Stokes (2002).[18] Craswell et al. (1995) speculate that larger companies are more likely to benefit from hiring an industry expert, and are also more capable of paying higher audit fees. This finding may also help to explain the lack of a fee premium for the 78 smaller than average companies in the sample that were audited by city industry leaders who were not also national leaders.

VI. DISCUSSION

The purpose of this study was to re-examine the effects of industry expertise on Big 5 audit fees using a new research design based on industry rankings to denote industry audit experts, and testing the effect of auditor industry expertise at both firm-wide view and office-level of analysis. Our results provide support for the office-level but not the firm-wide view of auditor industry expertise. Specifically, our tests show that being one of the top two firms nationally does not translate to a fee premium unless the auditor is also the city-specific industry leader. The national top two firms *do not* earn a premium in those cities where they are not the city leader. We also document that national industry leadership in 21 of the 24 industries in the study is driven almost exclusively by the individual offices in which the accounting firm is also the city-level industry leader. We further document that being the industry leader in multiple cities does not increase the fee premium relative to those cases where the Big 5 firm is the city leader in only one city.

In sum, the evidence does not support the existence of positive network externalities across offices, which is assumed to be the case in the firm-wide view of industry expertise. Rather, our tests show the opposite, namely, the industry expertise that is priced by the audit market is city-specific and a function of local-office industry leadership. The firm-wide and international views of Big 5 accounting firms are important in understanding their organization structures and operations. However, our findings provide evidence that the office-level perspective and city-specific audit markets are also important to fully understanding the operations of Big 5 accounting firms, and recent research has begun to investigate office-level behavior of Big 5 firms (e.g., Craswell et al. 2002; Reynolds and Francis

---

[18] The effect of the interaction term on the premium in Model (3) in Table 3 is estimated by calculating the fee premium for ±1 standard deviation of *LTA*, holding all other factors constant. In our sample the premium ranges from close to zero for an observation with minus 1 standard deviation of *LTA*, to a premium of 30 percent for an observation with plus 1 standard deviation of *LTA*.

2000). As a caveat, we do note that for 78 observations in 24 city-industry combinations, which represent 10 percent of total audit fees in the sample, the city-level industry leader does not earn a premium for industry expertise and this may be related to the relatively smaller size of companies in these particular city-industry combinations.

The empirical results in the study are necessarily limited to the Australian audit market, and with only five major cities it may be difficult to perfectly isolate firm-wide vs. office-level measures of auditor industry expertise. However, there are reasons to believe similar results would be found in the U.S. audit market. The reason is that the U.S. audit market is far more geographically dispersed than the Australian market. In our sample, 84 percent of Australian publicly listed companies are concentrated in just three cities: Melbourne, Perth, and Sydney. By contrast, Francis et al. (1999, 194) report that the 10 largest U.S. cities have only 50 percent of U.S. publicly listed companies. Big 5 practices are also more widely dispersed in the U.S. Francis et al. (1999, 195) document that among the Big 5 firms' 10 largest offices, there are only five common cities. This means there is greater geographical dispersion in the U.S. than in Australia of both clients and auditors, which means that it may be more difficult for Big 5 firms to create uniform firm-wide reputations for industry expertise and to achieve positive network externalities across all offices. In other words, with greater dispersion there is a greater potential for variations in Big 5 reputations and clienteles across cities, and thus it is possible that one would observe a significant local-office effect on the perception and pricing of industry expertise in the United States.

On the other hand, it may be easier for Big 5 accounting firms in Australia to develop unique office-specific reputations since there are only five primary cities in the country. In contrast, the fact that Big 5 firms have many more practice offices in the United States may make it more difficult for individual offices to develop unique reputations that differ markedly from the firm's overall national reputation. Because the U.S. offers a richer multi-office setting than Australia, an obvious direction for future research is to use the new U.S. fee disclosures to further examine the interplay of firm-wide and office-level characteristics of accounting firms, and the audit market's perception and pricing of auditor industry expertise.

REFERENCES

Balsam, S., J. Krishnan, and J. Yang. 2002. Auditor industry specialization and earnings quality. Working paper, Temple University.

Behn, B., J. Carcello, D. Hermanson, and R. Hermanson. 1997. The determinants of audit client satisfaction among clients of Big 6 firms. *Accounting Horizons* 11: 7–24.

Bental, B., and M. Spiegel. 1995. Network competition, product quality, and market coverage in the presence of network externalities. *Journal of Industrial Economics* 43: 197–208.

Caminal, R., and X. Vives. 1996. Why market shares matter: An information-based theory. *Rand Journal of Economics* 27 (Summer): 221–239.

Carcello, J., R. Hermanson, and N. McGrath. 1992. Audit quality attributes: The perceptions of audit partners, preparers, and financial statement users. *Auditing: A Journal of Practice & Theory* 11: 1–15.

Chan, D., A. Ferguson, D. Simunic, and D. Stokes. 2002. A spatial analysis and test of oligopolistic competition in the market for audit services. Working paper, University of British Columbia.

Chan, P., M. Ezzamel, and D. Gwilliam. 1993. Determinants of audit fees for quoted UK companies. *Journal of Business Finance and Accounting* 20: 765–786.

Craswell, A., J. Francis, and S. Taylor. 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics* 20: 297–322.

———, D. Stokes, and J. Laughton. 2002. Auditor independence and fee dependence. *Journal of Accounting and Economics* 33: 253–275.

DeFond, M., J. Francis, and T. Wong. 2000. Auditor industry specialization and market segmentation: Evidence from Hong Kong. *Auditing: A Journal of Practice & Theory* 19: 49–66.

Ferguson, A., and D. Stokes. 2002. Brand name audit pricing, industry specialisation and leadership premiums post Big 8 and Big 6 mergers. *Contemporary Accounting Research* 19: 77–100.

Francis, J., 1984. The effect of audit firm size on audit prices: A study of the Australian market. *Journal of Accounting and Economics* 6: 133–151.

———, and D. Stokes. 1986. Audit prices, product differentiation, and scale economies: further evidence from the Australian audit market. *Journal of Accounting Research* 24: 383–393.

———, and D. Simon. 1987. A test of audit pricing in the small-client segment of the U.S. audit market. *The Accounting Review* 62: 145–157.

———, D. Stokes, and D. Anderson. 1999. City markets as a unit of analysis in audit research and the re-examination of Big 6 market shares. *Abacus* 35: 185–206.

Gilson, R., and R. Mnookin. 1985. Sharing among the human capitalists: An economic inquiry into the corporate law firm and how partners split profits. *Stanford Law Review* 37: 313–392.

Hellofs, L., and R. Jacobsen. 1999. Market share and customers' perceptions of quality: When can firms grow their way to higher versus lower quality? *Journal of Marketing* 63 (January): 16–25.

Hogan, C., and D. Jeter. 1999. Industry specialization by auditors. *Auditing: A Journal of Practice & Theory* 18: 1–17.

Katz, M., and C. Shapiro. 1985. Network externalities, competition, and compatability. *American Economic Review* 75: 424–440.

Krishnan, G. 2003. Does Big 6 auditor industry expertise constrain earnings management? *Accounting Horizons* 17: forthcoming.

Narayanan, V. 1995. Moral hazard in repeated partnerships. *Contemporary Accounting Research* 11: 895–917.

Pearson, T., and G. Trompeter. 1994. Competition in the market for audit services: The effect of supplier concentration on audit fees. *Contemporary Accounting Research* 11: 115–135.

Penno, M., and B. Walther. 1996. The concentration of local markets: A study of accounting, advertising and law. *Accounting Horizons* 10: 88–99.

Reynolds, J. K., and J. Francis. 2000. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30: 375–400.

Smallwood, D., and J. Conlisk. 1979. Product quality in markets where consumers are imperfectly informed. *Quarterly Journal of Economics* 93: 1–23.

Solomon, I., M. Shields, and O. R. Whittington. 1999. What do industry specialists know? *Journal of Accounting Research* 37: 191–208.

Wallman, S. 1996. The future of accounting, Part 3: Reliability and auditor independence. *Accounting Horizons* 10: 76–97.

Zhou, J., and R. Elder. 2002. Audit firm size, industry specialization and earnings management by initial public offering firms. Working paper, Binghampton University and Syracuse University.

THE ACCOUNTING REVIEW
Vol. 78, No. 2
2003
pp. 449–469

Discretionary Risk Disclosures

Bjorn N. Jorgensen
Michael T. Kirschenheiter
Columbia University

ABSTRACT: We model managers' equilibrium strategies for voluntarily disclosing information about their firm's risk. We consider a multifirm setting in which the variance of each firm's future cash flow is uncertain. A manager can disclose, at a cost, this variance before offering the firm for sale in a competitive stock market with risk-averse investors. In our partial disclosure equilibrium, managers voluntarily disclose if their firm has a low variance of future cash flows, but withhold the information if their firm has highly variable future cash flows. We establish how the manager's discretionary risk disclosure affects the firm's share price, expected stock returns, and beta, within the framework of the Capital Asset Pricing Model. We show that whereas one manager's discretionary disclosure of his firm's risk does not affect other firms' share prices, it does affect the other firms' betas. Also, we demonstrate that a disclosing firm has lower risk premium and beta *ex post* than a nondisclosing firm. Finally, we show that *ex ante*, the expected risk premium and expected beta of each firm are higher under a mandatory risk disclosure regime than in the partial disclosure equilibrium that arises under a voluntary disclosure regime.

Keywords: *disclosures; risk; asset pricing.*

I. INTRODUCTION

Theoretical research on the disclosure of financial accounting information focuses on disclosures about financial statement items, such as asset values, earnings, operating cash flows, or management forecasts of earnings. This focus was natural as, until recently, virtually all mandated disclosures related to financial statement items. However, a recent change in financial disclosure requirements, Securities and Exchange Commission (SEC) Financial Reporting Release No. 48 (FRR No. 48), requires firms to disclose information about risk. We model the effect of additional disclosures of risk information beyond the disclosures required by FRR No. 48. Our study is the first attempt to model managers' voluntary risk disclosures.

Our research objective is to investigate how a manager chooses his disclosure strategy when he has discretion over whether he informs investors about the variance of his firm's

---

We appreciate the helpful comments and suggestions of Mary Ellen Carter, Raffi Indjejikian (the associate editor), Jennifer Kao, Irene Kim, Carolyn Levine, Russell Lundholm, Linda Vincent, two anonymous referees, and seminar participants at the 2002 American Accounting Association Annual Meeting, 11th Annual Financial Economics and Accounting Conference at the University of Michigan, 2000 Burton Workshop at Columbia University, Carnegie-Mellon University, Society for Economic Design 2000, University of Alberta, University of British Columbia, Universidad Carlos III, Baruch College of the City University of New York, Harvard University, University of Minnesota, and University of Rochester. We thank Kelly Puig for editorial comments.

Editor's note: This paper was accepted by Raffi J. Indjejikian, *Ad Hoc* Editor.

Submitted September 2001
Accepted November 2002

cash flows. In our model, each firm's future cash flows depend on two sources of risk: one market-wide, the other firm-specific. Since disclosure of sensitivity to market risk is mandated by FRR No. 48, we assume that investors know the variance of the market-wide cash flow factor. However, we assume that initially investors do not know the variance of firm-specific cash flows. Each risk-neutral manager can convey the information about his firm's firm-specific cash flow by disclosing, at a cost, his firm's cash flow variance before offering the firm for sale in a competitive market of risk-averse investors.

This paper contributes to the literature in two ways. First, we identify the equilibrium disclosure strategy in a general exchange economy where each manager has the discretion to issue a costly disclosure about his firm's risk. We show that managers of firms with low variance of future cash flows voluntarily disclose their (favorable) risk information, whereas managers of firms with high variance of future cash flows do not disclose.

This first contribution extends results on a manager's discretionary disclosure of earnings forecasts (Verrecchia 1983, 1990) to a manager's discretionary disclosure of *risk*. Analogous to the results in Verrecchia (1990), we identify factors that affect the manager's choice of the threshold level of the firm's cash flow variance, above which the manager will not make a risk disclosure. We show that, *ceteris paribus*, a firm that discloses risk has a higher share price than one that does not, but, relative to a voluntary disclosure regime, imposing mandatory full disclosure of firm risk will lower the *ex ante* share price of each firm. Firm value falls because mandating risk disclosure forces firms that would not disclose in a voluntary regime to incur disclosure costs.

Our second contribution stems from analyzing multiple firms and relating their risk disclosure decisions to stock returns. We establish that even when some of the managers choose not to disclose, the expected returns behave in a manner consistent with the Capital Asset Pricing Model (CAPM). Under the CAPM, a firm's risk premium is the expected excess return on its stock over the riskless bond, which is proportional to the risk premium on the market, that is, firm risk premium = beta × market risk premium. This allows us to separate the effect of investors' uncertainty about firms' risk premia into two components. The first, direct effect, from Barry and Brown (1985), is that an increase in investors' uncertainty about the variance of the firms' future cash flows leads to a higher risk premium. In our setting, a second, indirect effect emerges. Because greater investor uncertainty increases the range of variances under which a manager will make voluntary disclosures about his firm's risk, this second indirect effect lowers the risk premium. When the latter indirect effect dominates the former direct effect, greater investor uncertainty about the firm's variance of future cash flows will *decrease* the risk premium. Which effect dominates remains an empirical question.

Our second contribution also enables us to investigate how a manager's voluntary disclosure about his firm's risk affects the CAPM parameters of his own and other firms. First, we find that a firm's beta and risk premium both increase if the manager discloses a high variance, and they increase further in the absence of disclosure. Second, we find that the risk premium of each firm's stock return depends only on that firm's own risk disclosures. Third, we specify how the beta of each firm and the risk premium on the return on the market portfolio depend on whether other firm managers disclose and what variance they disclose. In particular, we show that disclosure by one manager will raise the betas of other firms in the economy, provided the firms have positive betas, while leaving the share prices of these other firms unaffected.

Recent empirical research concludes that FRR No. 48 disclosures are useful to investors (e.g., Rajgopal 1999; Linsmeier et al. 2002). Our model provides a theoretical framework for how fully rational investors may use voluntary disclosures about the firm-specific variance of future firm cash flows, which are beyond the risk disclosures required under FRR No. 48. Our findings provide several empirically testable hypotheses. For example, our results suggest that firms are more likely to voluntarily disclose information about risk in industries where investors are more uncertain about the level of firm-specific risk, such as nondiversified research and development intensive firms. Our work also suggests that a firm will have a lower risk premium and a lower beta if its manager discloses risk information than if the manager does not disclose. In addition, we predict a lower market risk premium and a higher beta for a firm if other firms disclose, but predict that a firm's risk premium remains unaffected by the disclosure of other firms. Further, we predict that expected betas will increase if regulators expand disclosure requirements beyond FRR No. 48 to require new disclosures of firm-specific variance. Collectively, our results suggest that when conducting or interpreting results from event studies, researchers should recognize that abnormal returns based on the market model may be biased if betas have changed in response to risk disclosures by the firm itself or by other firms, or in response to changes in disclosure requirements.

II. INSTITUTIONAL BACKGROUND AND RELATED RESEARCH

Risk Disclosure and the Financial Reporting Environment

Risk disclosures are an increasingly important part of financial reporting. Significant losses from firms' use of financial instruments in the 1990s prompted investors, regulators, auditors, and other private bodies to call for expanded disclosures on risks of financial instruments.[1] The SEC concluded that financial disclosures relating to risk were often incomplete and, in some circumstances, misleading. To address this problem, in January 1997, the SEC issued Financial Reporting Release No. 48, *Disclosure of Accounting Policies for Derivative Financial Instruments and Derivative Commodity Instruments and Disclosure of Quantitative Information About Market Risk Inherent in Derivative Financial Instruments, Other Financial Instruments and Derivative Commodity Instruments* (FRR No. 48). While not the only regulation to require disclosure of risk-related measures, FRR No. 48 is the current standard for risk-related disclosure.[2]

FRR No. 48 requires firms to disclose quantitative and qualitative information about market risk exposures. Jorion's (2002) empirical evidence that the variance of banks' trading revenues is related to their risk disclosures suggests that these disclosures are informative about the firm's risk.[3] Consequently, we assume that the FRR No. 48 disclosures provide information about the sensitivity of the firm's cash flows to market-wide risk factors, so that this information is common knowledge. FRR No. 48 disclosure requirements pertain only to market risk related to financial instruments and derivatives. As a consequence, investors may not know the sensitivity of the firm's other cash flows to market-wide risk. Despite this caveat, we believe that FRR No. 48 disclosures suggest that investors are more likely to have common knowledge about market risk than about other firm-specific risk. Hence, we assume that investors use the voluntary disclosures to make inferences concerning firm-specific risk.

---

[1] Included in this list calling for additional disclosures were the American Institute of Certified Public Accountant (AICPA), the Association for Investment Management and Research (AIMR), and the Financial Executives Institute (FEI), see Linsmeier and Pearson (1997) for a survey discussing the losses from financial instruments and the calls for additional disclosures.

[2] Issued in June 1998, SFAS No. 133, *Accounting for Derivative Instruments and Hedging Activities*, encourages, but does not require, firms to disclose quantitative information about market risks of financial instruments (para. 15D). In September 1998, the U.K. Accounting Standards Board issued Financial Reporting Standard No. 13 about risk disclosure. Also, Bank of International Settlements regulates financial institutions worldwide based on Value-at-Risk, which is one of the types of quantitative risk disclosures under FRR No. 48.

[3] Moreover, while one may question whether the FRR No. 48 disclosures effectively communicate market-wide risk information, Hodder et al. (2001) generally concede that financial statement users interpret these disclosures as doing so.

Risk-Disclosure-Related Research

To clarify how our study of managers' voluntary risk disclosures contributes to the literature, we review the existing research and explain why these models fail to address our research question.

First, Barry and Brown (1985) and Coles et al. (1995), among others, investigate the effect of investors' uncertainty about the variance of cash flows on market values, stock returns, and betas. They establish that a firm's beta increases in the investors' uncertainty about the variance of a firm's cash flows. Neither of these papers, however, allows firm managers to disclose the variance of their firms' cash flows, whereas managers' risk disclosure decisions and the effect of these decisions on share prices are the focus of our study.

Second, recent accounting studies explore the effect of the managers' disclosure choices on the precision of earnings. For example, Dye (1990) and Penno (1996) investigate a manager who chooses the precision of the earnings reported, and Kirschenheiter and Melumad (2002) consider a setting where managers' accrual decisions affect investors' inferences about the precision of earnings. These papers do not allow the manager to either disclose or refrain from disclosing the precision. Hence, none of these studies consider the effect of a manager's discretionary risk disclosure.

Third, recent research on hedge accounting, including DeMarzo and Duffie (1995), Melumad et al. (1999), and Kanodia et al. (2000), evaluates managers' optimal hedging decisions when each manager must disclose his firm's risk relative to the benchmark where the manager cannot disclose his firm's risk. Because these studies do not allow the manager to choose whether to report or withhold his private information regarding the firm's risk exposure, they do not bear on discretionary risk disclosure. In contrast, we evaluate mandatory disclosure of risk relative to our alternative benchmark of voluntary disclosures, while abstracting from the manager's decision to hedge.

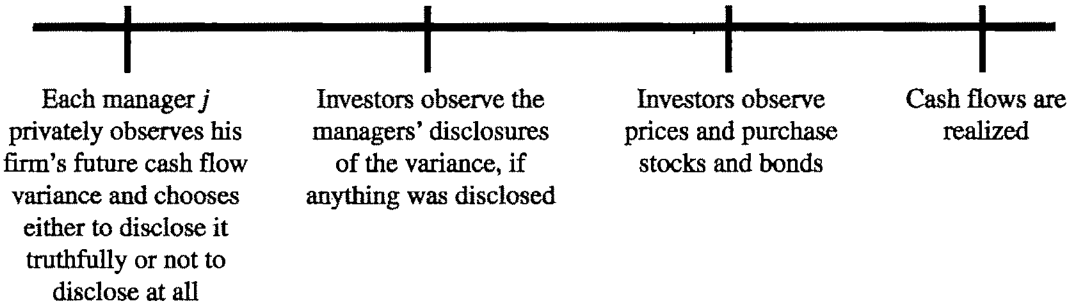
In summary, no model currently exists that explains how a manager decides whether to voluntarily disclose the variance of his firm's future cash flow that he privately observes. To address this question, we first derive the share prices in a general exchange economy with risk-averse investors where managers can make such a disclosure choice. We then use these share prices to develop the CAPM and investigate the relation between the CAPM parameters (e.g., the beta and the excess return on the market) and the managers' disclosure decisions.

III. MODEL

Consider the following single period model of a capital market with I investors who trade in J firms' risky stock and a riskless bond. At time 0, each manager privately observes the variance of his firm's risky cash flows. The managers then simultaneously decide whether to truthfully disclose their firm's variance to all investors or to refrain from disclosing. If the manager of firm j discloses, then he incurs a disclosure cost of $C_j > 0$, which reduces the firm's cash flows at time 2. At time 1, investors trade in the J risky stocks and the riskless bond after observing what disclosure, if any, each manager made. At time 2, investors consume and the economy ends. Figure 1 illustrates the timeline. The following detailed assumptions are common knowledge to all investors and managers.

> **Assumption A1 (Assumption on Cash Flows):** The exchange economy includes J firms indexed by $j = 1,...,J$. Firm j has a risky investment project in place that pays X_j at time 2. The lower case, x_j, denotes the realization of the random variable X_j. We assume a single market-wide factor model describes cash flow uncertainty, that is, $X_j = \mu_j + \gamma_j F + \varepsilon_j$, where μ_j is the expected cash flow, F denotes the market-wide cash flow factor, γ_j is the factor loading, or sensitivity of firm j's cash flows to the market-wide factor, and the firm-specific cash flow is ε_j for $j = 1,...,J$. Since the randomness in the cash flow of firm j depends only on these two factors, we refer to F and ε_j as the market-wide risk factor and firm-specific risk factor for firm j, respectively. The total supply of the jth risky asset is one share, with share price of P_j.
>
> **Assumption A2 (Assumptions on Distributions):** Assume the market-wide factor F is distributed $N(0, \sigma^2)$, and firm-specific cash flow is distributed $\varepsilon_j | v_j \sim N(0, v_j)$, where the variance v_j is the realization of a continuous, non-negative random variable V_j and is drawn from a probability distribution $f_j(v_j)$. Hence, the conditional distribution of a firm's future cash flows is $X_j | v_j \sim N(\mu_j, \gamma_j^2\sigma^2 + v_j)$.
>
> **Assumption A3 (Assumption on Managers):** At time 0, firm j's manager privately observes the realized variance, v_j. He then chooses whether to truthfully disclose v_j, incurring cost $C_j > 0$ if he does disclose and no cost if he does not disclose.[4] Manager j's disclosure strategy is characterized by $N_j \subseteq \Re^+$, the set of values of the variance that the manager j will not disclose. Let $\underline{N}$ denote the vector of disclosure strategies for all J managers and let $\underline{N}_{-j}$ denote the $(J - 1)$-dimensional vector of all firms' disclosure strategies except firm j. Each manager holds the same beliefs about other managers' disclosure strategies, as do the investors.
>
> **Assumption A4 (Assumption on Investors):** There are I individuals in the market indexed by $i = 1,...,I$ who have constant absolute risk aversion, $a_i > 0$. Each investor

FIGURE 1
Timeline


Each manager j privately observes his firm's future cash flow variance and chooses either to disclose it truthfully or not to disclose at all
Investors observe the managers' disclosures of the variance, if anything was disclosed
Investors observe prices and purchase stocks and bonds
Cash flows are realized

---

[4] We assume that the disclosure costs are exogenous, and we do not model the investment decision. Endogenizing the disclosure costs, as in Gigler (1994) among others, and explicitly modeling the managers' investment decisions would represent an extension of our model. Also, while we assume any disclosure the manager makes is entirely truthful, Stocken's (2000) arguments could support the truthful risk disclosures in our setting.

allocates his initial wealth of W_i^0 between purchasing S_{ij} shares of the jth stock and investing B_i in riskless bonds. Investors take as given the share price of each firm, P_j, and the return on the bond, R_f. Let $\hat{N}_j \subseteq \Re^+$ denote the set of values of the variance in the firm's future cash flows that the investors believe is manager j's nondisclosure set, where the circumflex (or "hat") over the N_j denotes investors' inferences concerning the disclosure strategy that the manager has adopted rather than the manager's actual disclosure strategy. Let $\hat{\underline{N}}$ denote the J-dimensional vector of these inferred disclosure strategies. If investors expect full disclosure but a manager does not disclose, then investors infer the worst; that is, the investors believe that the manager observed but did not disclose the highest possible variance. Finally, consistent with A1 and A3, let $P_j = P_j(v_j, N_j|\underline{N}_{-j}, \hat{\underline{N}})$ denote price of firm j's shares.[5]

We maintain Assumptions A1–A4 throughout the paper. In particular, A2 denotes the marginal effect of the single market-wide factor, F, on firm j's cash flows, as γ_j. Since FRR No. 48 requires that firms disclose their exposures to market risk, we assume that each γ_j is common knowledge. In addition, we assume that the variance in underlying risk factors is observable. Hence, we assume that the variance of the market-wide factor, σ^2, is also common knowledge. Assumption A3 specifies that the manager observes the variance of the firm-specific cash flow risk, but the investors do not. Thus, we assume that the manager must decide whether to disclose the variance of the firm-specific cash flows. Firm j's disclosure of the variance of total future cash flows then becomes equivalent to the disclosure of firm-specific variance, v_j.

Building upon the beliefs and strategies of the managers and investors discussed above, we next specify their payoffs and define our equilibrium concept. Investor i allocates his initial wealth, W_i^0, between the riskless bond and the risky stocks. Each investor trades the bond and the shares in the risky assets to maximize his terminal wealth, W_i, given his information about the variance of the asset's return. Using $1_{\{v_j \notin N_j\}}$ as the indicator function if the manager discloses, we have:

$$W_i = B_i R_f + \sum_{j=1}^{J} S_{ij}\,(X_j - 1_{\{v_j \notin N_j\}} C_j).$$

Hence, each investor's maximization problem is:

$$\max_{\{S_{ij}\}_{j=1}^{J},\, B_i} \quad E\lfloor -e^{-a_i W_i} \rfloor$$

subject to:

$$B_i + \sum_{j=1}^{J} S_{ij} P_j \leq W_i^0$$

where the investor takes the prices of all assets as given. For each firm j, the price clears the market for shares; that is, $\Sigma_{i=1}^{I} S_{ij} = 1$. The equilibrium share price of each firm, P_j, is set by the investors' demand for the firm's stock, which in turn depends on what, if anything, the manager of firm j disclosed and how investors interpret no disclosure.

---

[5] This is abbreviated notation because, in equilibrium, the share price of firm j depends on whether its manager discloses. Only when the manager of firm j discloses do investors observe the variance and condition the share price on the realized value of the variance, v_j. Further, while we allow for the possibility that firm value depends on whether other managers disclose and what disclosures they make, the disclosure made by competing firms do not affect the equilibrium price of firm j.

A *credible* equilibrium to this game is defined as follows:

a) Each manager chooses his disclosure strategy to maximize the value of his firm j, net of any disclosure cost, given investors' inferences about the managers' disclosure strategies. That is, $P_j(v_j, N_j|\underline{N}_{-j}, \underline{\hat{N}}) \geq P_j(v_j, N_j'|\underline{N}_{-j}, \underline{\hat{N}})$ for each j, each $v_j \geq 0$ and for all $N_j' \subseteq \Re^+$.
b) In equilibrium, $\underline{N} = \underline{\hat{N}}$ (investors correctly anticipate the managers' disclosure strategies).

We refer to the endogenous vector of disclosure strategies, $\underline{N}$, that simultaneously solve both Conditions (a) and (b) as a credible equilibrium.

IV. CONSEQUENCES OF RISK DISCLOSURES

We present our results in two subsections. First, we establish how a manager's optimal risk disclosures affect his firm's share prices. Second, we establish how a manager's risk disclosures affect his firm's and other firms' stock returns. All proofs appear in the Appendix.

Effect of Managers' Risk Disclosures on Share Prices

We begin by characterizing the disclosure strategies that arise in a credible voluntary disclosure equilibrium in Lemma 1:

Lemma 1: If equilibrium share prices are monotonically decreasing in the disclosed variance of the firm's cash flows, then the nondisclosure region is determined by a threshold level of variance where the manager discloses only if he (privately) observes that the variance of his firm's future cash flow is below this level (i.e., the nondisclosure region is upper-tailed).

Lemma 1 states that if the equilibrium share prices decrease in the disclosed variance of the firm's cash flows, then managers withhold unfavorable news and disclose favorable news about this variance.[6] Hence, Lemma 1 allows us to express firm j's disclosure strategy, N_j, by its disclosure threshold, y_j.[7]

Lemma 1 describes the credible disclosure strategy based on the presumption that share prices decrease in variances. Next, in Lemma 2, we show that share prices in this economy exhibit this property whether the manager discloses or withholds his private information about the variance of his firm-specific cash flows.

Lemma 2: Let $a = (\Sigma_{i=1}^{I} a_i^{-1})^{-1}$ measure investors' aggregate risk aversion and let $\overline{\gamma} = \Sigma_{j=1}^{J} \gamma_j$ be the sum of the factor loadings. Given any vector of inferred disclosure strategies, $\underline{\hat{N}}$, the share price of firm j is given as follows:

---

[6] Lemma 1 generalizes to all types of disclosures, not just risk disclosures. We are grateful to an anonymous referee for suggesting this generalization.

[7] Lemma 1 shows that credible disclosure strategies will be one of three types: no disclosure, full disclosure, or partial disclosure strategies, where the manager discloses the variance if it is less than some threshold level, but does not disclose his firm's variance if it exceeds this threshold. To focus on how risk disclosure regulation affects the relation between the manager's disclosure strategy and his firm's share price and the firm's cost of capital, the discussion in the body of the paper concentrates on partial disclosure equilibria.

a) If manager j discloses his private information about the variance in the firm-specific cash flow of his firm ($v_j \notin N_j$), then the price of his firm is:

$$P_j(v_j, N_j|\underline{N}_{-j}, \hat{\underline{N}}) = \{\mu_j - C_j - a(\gamma_j\sigma^2\bar{\gamma} + v_j)\}R_f^{-1}.$$

b) If manager j withholds his private information about the variance in the firm-specific cash flow of his firm ($v_j \in N_j$), then there exists a scalar, $\delta_j(\hat{N}_j)$, such that the price of his firm is:

$$P_j(v_j, N_j|\underline{N}_{-j}, \hat{\underline{N}}) = \{\mu_j - a(\gamma_j\sigma^2\bar{\gamma} + \delta_j(\hat{N}_j)E[V_j|v_j \in \hat{N}_j])\}R_f^{-1}.$$

Four aspects of Lemma 2 are noteworthy. First, the price of the firm given disclosure decreases in the variance of the firm's cash flows. Consequently, Lemma 1 implies that we can describe each manager's nondisclosure set by a disclosure threshold for each firm j, y_j, such that the manager does not disclose the observed variance, v_j, if and only if $v_j \geq y_j$. We therefore can write part (b) of Lemma 2 as:

$$P_j(v_j, N_j|\underline{N}_{-j}, \hat{\underline{N}}) = \{\mu_j - a(\gamma_j\sigma^2\bar{\gamma} + \delta_j(y_j)E[V_j|V_j \geq y_j])\}R_f^{-1}.$$

Second, comparing the price equations (a) and (b) of Lemma 2 reveals similarities. When the manager discloses the variance, share prices separate the effect of the mean and the variance of total cash flows, with the variance being multiplied by the aggregate risk aversion, a, as in Wilson (1968). When the manager discloses, the total cash flow variance equals the sum of: (1) the market-wide risk portion, $\gamma_j\sigma^2\bar{\gamma}$, plus (2) the disclosed variance, v_j, of the firm-specific cash flow term. When the manager refrains from disclosing the firm-specific cash flow variance, the product $H(y_j) = \delta(y_j)E[V_j|V_j \geq y_j]$ replaces the variance of the firm-specific cash flow term in the price equation. This product consists of the expected variance in the firm's cash flows given no disclosure, $E[V_j|V_j \geq y_j]$, multiplied by the new risk adjustment factor $\delta_j(y_j) \geq 1$. This scalar risk adjustment factor exceeds unity because no disclosure implies additional uncertainty, so that investors in turn require a lower price.

Third, the size of the economy affects the pricing of the firm's shares. In the limiting case where there are infinitely many firms in the economy, the effect of firm-specific cash flow risk becomes negligible and does not affect share prices. When that happens, the role of discretionary disclosure of firm-specific variance disappears. However, our analysis applies to large, but finite, economies. This leads to our fourth and final point.

Lemma 2 shows us that the market sets share prices as if it adds the effects of the market-wide and firm-specific risks together. The manager chooses whether to disclose based on the difference in the share price when he discloses the variance vs. when he does not. Since the market-wide risk factor enters share price additively whether or not the manager discloses, it does not affect his disclosure decision. This fact proves useful in analyzing the manager's equilibrium disclosure strategy.

We now present the existence result regarding the credible partial disclosure equilibrium in Theorem 1.

Theorem 1: For some intermediate level of disclosure costs there exists a credible partial disclosure equilibrium.

Theorem 1 shows that, given Assumptions A1–A4, a partial disclosure equilibrium exists. However, Theorem 1 does not ensure uniqueness. To rule out an infinite multiplicity of partial disclosure equilibria, we introduce the following additional assumption.

> **Assumption A5 (Assumption of a Non-Constant Risk Adjustment):** The risk adjustment associated with nondisclosure, $H(y_j) \equiv \delta(y_j)E\lfloor V_j | V_j \geq y_j \rfloor$, is differentiable and its derivative differs from one for all y_j in the support of V_j, almost surely.

In our model, Assumption A5 does not hold trivially because it fails to hold when investors believe the variance has an exponential distribution. In that case, there exists a single disclosure cost level that supports a partial disclosure equilibrium and further, every threshold level supports an equilibrium at that cost. We therefore require Assumption A5 to perform comparative statics on the partial disclosure equilibrium.

Throughout the remainder of the paper, we maintain Assumptions A1–A5 and intermediate disclosure cost levels for all firms. These assumptions enable us to use the credible equilibrium of Theorem 1 as our voluntary disclosure benchmark. To gain further intuition regarding the managers' voluntary disclosure decision in the partial disclosure equilibrium, we present comparative statics for two examples of a single-firm economy. (Proofs for these two examples are available from the authors.)

> **Example 1:** Assume that investors' prior beliefs are that the cash flow variance is uniformly distributed. Then partial disclosure is the unique credible equilibrium for sufficiently low disclosure costs, and the resulting disclosure threshold is increasing in the mean and increasing in the standard deviation of investors' prior distributions over the variances. Further, an increase in the standard deviation of investors' prior distribution over the variance increases the probability that the manager discloses and decreases the firm's share price if the manager does not disclose.

This example illustrates the effect of changes in investors' prior distribution regarding the variance of the firm's cash flows. Increasing the standard deviation of the investors' prior beliefs about the variance increases the equilibrium disclosure threshold, y, and this in turn implies that disclosure is more likely. As discussed further below (see Corollary 3), this result extends earlier results in Barry and Brown (1985).

> **Example 2:** Assume that investors believe that the cash flow variance is Gamma distributed with parameters 2 and λ, and that investors' aggregate risk aversion is low such that $a^2 < 2\lambda$. Then partial disclosure arises as an equilibrium for intermediate disclosure costs. Furthermore, the associated disclosure threshold is decreasing in the disclosure costs, increasing in investors' risk aversion, and decreasing in λ.

The comparative static result on disclosure costs is qualitatively similar to the comparative static results reported for discretionary disclosures of management forecasts in Verrecchia (1990): higher disclosure costs lead to a higher disclosure threshold or lower likelihood of disclosure. We also demonstrate that an increase in investors' risk aversion leads to a higher disclosure threshold or higher likelihood of disclosure. To interpret the last comparative static result in Example 2, recall that increasing λ decreases both the mean and the standard deviation of investors' prior beliefs about the variance, V.[8] Consistent with Verrecchia (1990), we measure initial information asymmetry between investors and the manager by the standard deviation of investors' prior beliefs about the variance. Decreasing the standard deviation of investors' prior beliefs reduces the pressure from investors for disclosure and, hence, a lower disclosure threshold. Also, in this example, the partial disclosure threshold is chosen relative to the mean, so decreasing the mean decreases the threshold. Simultaneous decreases in both the mean and the standard deviation of investors' prior beliefs about the variance of the firm's future cash flows lead to a lower disclosure threshold, again similar to Verrecchia (1990).[9]

---

[8] This is so because the prior mean and standard deviations are $(2/\lambda)$ and $(\sqrt{2}/\lambda)$, respectively.

We next compare the effect of voluntary risk disclosure on share prices from two different perspectives. First, in a partial disclosure equilibrium, we establish how discretionary risk disclosure affects share prices. This is an *ex post* comparison of the share price after the variance has been realized. Second, we compare a firm's *ex ante* share price in a voluntary risk disclosure regime to its *ex ante* share price in a mandatory risk disclosure regime.

Corollary 1: In the partial disclosure equilibrium, a firm has higher share price if the manager discloses than if he does not.

Corollary 1 holds because disclosing firms have more favorable news. Corollary 1 adopts an *ex post* perspective, so we cannot use this result to infer how different disclosure decisions affect the share price of a single firm. To do this, we next adopt an *ex ante* perspective and address the question of which disclosure regime, mandatory or voluntary, produces higher share price for each firm. This comparison is given in the following corollary.

Corollary 2: Consider the benchmark of a voluntary risk disclosure regime. First, imposing a mandatory full risk disclosure regime will lower the *ex ante* expected share price of the firm. Second, imposing a mandatory no-disclosure regime will increase the *ex ante* expected share price of the firm.

The intuition behind Corollary 2 is as follows. First, under voluntary disclosure, the manager will choose not to disclose for some realizations of the variance. This implies that no disclosure results in higher firm value for those realizations than does disclosing. Hence, imposing mandatory risk disclosures reduces the *ex post* share prices for nondisclosers. Second, the result comparing voluntary disclosure to forbidden disclosure is similar to Verrecchia's (1983) result on disclosures of a signal about the firm's liquidating value. *A priori*, both managers and investors base their expected variance on their prior beliefs about variances, regardless of the regime in effect. Mandating no disclosure leaves the expected variance unchanged, but ensures that the firm will not incur the disclosure costs, which, in turn, increases share prices. Corollary 2 contributes to the literature by comparing voluntary to mandatory risk disclosure regimes in a general exchange economy with risk-averse investors and endogenous share prices.

Effect of Managers' Risk Disclosures on Market Returns

The equilibrium prices derived in connection with Theorem 1 enable us to extend the literature further by analyzing the CAPM in a partial disclosure equilibrium. Our investigation of the relation between risk disclosure decisions and the CAPM requires us to specify a return measure. We define returns using a gross measure for the stocks between time 1 and time 2, that is, the return on the stock of firm j is:

---

[9] We mention for completeness that the manager never discloses if disclosure costs are sufficiently high. Further, full disclosure can arise in Example 2, and this is the only equilibrium when the disclosure costs are sufficiently low.

$$R_j = \frac{(X_j - 1_{\{v_j \notin N_j\}} C_j)}{P_j(v_j, N_j | \underline{N}_{-j}, \underline{\hat{N}})}.$$

Similarly, we denote the gross return on the market portfolio as $R_m = W_m / W_m^0$, where W_m^0 and W_m are the initial and ending wealth in the economy, respectively. In our next result, we verify that the CAPM holds even when investors are uncertain about the firm-specific variance, and a manager's disclosure could resolve this uncertainty.

Theorem 2: The CAPM holds in each credible partial disclosure equilibrium, that is, the risk premium for each firm j can be expressed as:

$$E[R_j] - R_f = \beta_j(E[R_m] - R_f)$$

$$\text{where} \quad \beta_j = \frac{COV[R_j, R_m]}{VAR[R_m]}.$$

Theorem 2 enables us to analyze the effect of discretionary risk disclosure on CAPM parameters. To facilitate this presentation, we introduce one final assumption.

Assumption A6 (Assumption of Positive Share Prices): The following are positive: the share price of firm j under mandatory disclosure, the factor loading for firm j, or $\gamma_j > 0$, and the sum of the factor loadings, or $\bar{\gamma} > 0$.

Besides ensuring that the price of firm j is positive for each observable variance, Assumption A6 guarantees that the market risk premium is positive, that is, $E[R_m] - R_f > 0$. We next establish how a manager's disclosure decision affects the market risk premium and his own firm's CAPM beta and risk premium.

Corollary 3: Assume that manager j discloses the variance v_j. The market risk premium increases in the disclosed variance, that is, $\frac{\partial}{\partial v_j}(E[R_m] - R_f) > 0$. The higher the variance disclosed by the manager j, the higher the beta of firm j and the higher the risk premium of firm j, that is, $\frac{\partial \beta_j}{\partial v_j} > 0$ and $\frac{\partial}{\partial v_j}(E[R_j] - R_f) > 0$, respectively.

First, since the total market risk increases with the variance of firm j, the market premium must also increase. Second, given disclosure by manager j, that firm's beta increases with the disclosed variance because a higher disclosed variance implies that investors know that firm j's cash flow contributes a higher fraction to the total risk in the market. Mathematically, the risk premium of firm j equals the product of the market risk premium and firm j's beta. Hence, combining these two effects, the risk premium of firm j must also increase in the disclosed variance. It also follows that, *ceteris paribus*, a firm has lower beta and lower risk premium if its manager discloses risk than if its manager withholds the risk information.

Corollary 3 also suggests an alternative hypothesis regarding the relation between the initial level of investor uncertainty about the variance in a firm's cash flows and that firm's beta. Barry and Brown (1985) show that higher dispersion in investors' prior beliefs about the variance of the firm's future cash flow leads to a higher risk premium. In our setting, however, higher dispersion in prior beliefs has a second indirect effect: higher dispersion in investors' prior beliefs about the variance could lead to a (disproportionately) higher disclosure threshold (i.e., managers make more frequent risk disclosures). Thus, managers of firms whose investors face excessive estimation uncertainty may disclose so much more frequently that the net effect of greater investor prior uncertainty about the firm's risk premium actually *decreases* beta—the opposite to Barry and Brown's (1985) prediction.

Corollary 3 shows that manager *j*'s disclosure decision can affect his firm's beta and risk premium. We now consider whether other managers' disclosure decisions affect the beta and the risk premium of firm *j*. We answer these questions in the following Corollary.

Corollary 4: Assume $\gamma_k > 0$ and that manager k discloses his firm-specific variance, v_k. The beta of all other firms is decreasing in v_k, or $\frac{\partial \beta_j}{\partial v_k} < 0$ for $k \neq j$, but the risk premium of these other firms is unaffected by manager k's disclosure, or $\frac{\partial}{\partial v_k}(E[R_j] - R_f) = 0$ for $k \neq j$.

An increase in the variance disclosed by manager *k* decreases the beta of firm *j* because the fraction of the total market risk contributed by firm *k* has increased. This further implies that, *ceteris paribus*, firm *j*'s beta is lower when manager *k* does not disclose (because firm *k*'s variance is so high that it is in the nondisclosure region) than when manager *k* discloses.

However, manager *k*'s disclosures do not affect the risk premium of firm *j*. Recall that the risk premium of firm *j* equals the product of the market risk premium and firm *j*'s beta. The decrease in the beta of firm *j* caused by the disclosure by firm *k* of a relatively high variance, v_k, offsets the increase in the market risk premium caused by the increased risk, leaving the product of the market risk premium and the beta of firm *j* unchanged. The share prices in Lemma 2 provide the intuition for this result. The share price of firm *j* does not depend on the disclosures made by other managers, so neither does the expected return of firm *j*. Consequently, a manager's risk disclosure decision does not affect the risk premium of the other firms in the economy. In summary, our results suggest that, in our model, one firm's risk disclosures affect the betas, but not the share prices or risk premia, of the other firms in the economy.

Building on the analysis of the effect of managers' voluntary risk disclosure decisions on CAPM parameters, we next compare these parameters across risk disclosure regimes. More specifically, we compare the *ex ante* expected beta of firms in a voluntary risk disclosure regime to the *ex ante* expected beta of those same firms if they were required to make risk disclosures.

Corollary 5: *Ex ante*, before the firm's cash flow variance is realized, the expected risk premium and the expected beta are higher under mandatory risk disclosure than under a voluntary risk disclosure regime. Further, the expected risk premium and the expected beta are higher under a voluntary risk disclosure regime than under a no-disclosure regime.

Corollary 5 states that the expected risk premium for each firm will be higher under mandatory full disclosure of firm risk than under either voluntary risk disclosures or mandated nondisclosure. Thus, a firm's cost of capital, as measured by its expected risk premium, will be weakly higher in a mandatory disclosure regime than in a voluntary risk disclosure regime. To understand the intuition for this result, consider what managers cannot do when disclosure becomes mandatory. In a voluntary regime, managers with unfavorable news can choose to withhold their private information. Investors incur a cost in terms of an additional risk premium, but this cost is less than investors' direct costs if the manager had been forced to disclose. Under mandatory full disclosure of firm risk, managers must incur the higher cost of disclosing unfavorable news, driving share prices down and driving firm risk premia up in equilibrium. Moving from a voluntary to a mandatory regime appears to redistribute wealth from current to future shareholders, but the net welfare effect will depend on the level and the nature of the disclosure costs incurred and any other unmodeled potential benefits of disclosure.

V. SUMMARY

Recent regulation requiring firms to disclose the sensitivity of certain of their future cash flows to market-wide risks highlights the need for a theoretical analysis of managers' risk disclosure decisions. Our research objective is to describe a manager's risk disclosure strategies and to assess the possible effects of requiring disclosure of a second aspect of risk—firm-specific risk. We consider a multifirm setting in which the manager of each firm privately observes the variance of his firm's future cash flows, and then maximizes firm value by choosing to either disclose truthfully at a cost or else refrain from disclosing. After observing all managers' discretionary disclosure decisions, risk-averse investors trade, and their demand for shares determines the share price of each firm.

We endogenously derive the price of each firm's shares. We show that these prices depend on market-wide risk, but that the manager's disclosure decision does not. Hence, in our setting, each manager's discretionary risk disclosure decision rests solely on the firm-specific cash flow risk his firm faces and is independent of other managers' disclosure decisions. From an *ex post* perspective, we find that in a voluntary risk disclosure regime, firms have higher share prices if their managers disclose the firm's risk than if they do not. From an *ex ante* perspective, however, imposing mandatory risk disclosure requirements lowers the firms' share prices relative to share prices in a discretionary disclosure regime. Requiring firms to make risk disclosures that they would not make voluntarily forces firms to incur additional disclosure costs that decrease firm value.

We verify that the Capital Asset Pricing Model holds in our setting when managers use a partial disclosure strategy, and we establish how a manager's disclosure strategy choice affects his firm's beta and risk premium. For example, we find that a firm's beta, its risk premium, and the market risk premium are all higher if a firm fails to disclose the variance of its firm-specific cash flows. Also, one firm's beta is higher if other firms' managers disclose than if they do not. In addition, a firm's expected risk premium and its expected beta are both higher in a mandatory risk disclosure regime than in a voluntary disclosure regime.

Our results offer empirically testable hypotheses. First, when investors are more uncertain about the level of the variance in the firm's future cash flows prior to disclosure, we predict that voluntary risk disclosure is more likely. Second, for a firm with positive beta, we predict that the risk premium on its stock, its beta, and the risk premium on the market increase as the disclosed variance increases, and increase further under nondisclosure. Third, for firms with positive betas, we predict that a higher variance disclosed by one firm lowers the betas for other firms. Finally, we predict that an increase in the mandatory disclosure requirements will increase the firms' average beta and risk premium.

Our analysis has a number of limitations, all of which we view as raising questions for future research. First, we presume that the disclosures mandated by FRR No. 48 ensure that managers fully disclose the sensitivity of their firms' cash flows to market-wide risk factors. However, FRR No. 48 requires firms to disclose the sensitivity to market risk factors of cash flows related only to financial instruments and derivatives. Thus, a question for future research is: Does the manager's discretionary disclosure decision change if investors do not know the sensitivity of all of the firm's cash flows to market-wide risk? Second, while FRR No. 48 mandates disclosure, it allows for a choice among three formats. How do managers choose among these formats? Third, the benefits from disclosure are limited in our setting because we do not consider the possible interaction between risk disclosure and managers' real investment decisions. How might a manager's discretionary disclosures about his firm's risk affect the firm's investment decisions? Fourth, we assume that investors are price takers. If we assume some investors are informed traders, does the manager's risk disclosure affect total order flows and bid-ask spreads? Fifth, we do not consider other types of disclosures. Do a manager's discretionary disclosures of earnings forecasts interact with his decision to disclose about risk? All of these questions remain open and deserve further exploration.

APPENDIX

Proof of Lemma 1

Manager *j*'s discretionary disclosure decision is characterized by $P_j(\hat{N}_j) \leq P_j(v_j)$ for all $v_j \notin N_j$ and $P_j(\hat{N}_j) \geq P_j(v_j)$ for all $v_j \in N_j$. It suffices to show that Lemma 1 holds for each firm individually, so we drop the *j* subscript and consider one manager's disclosure decision. The remainder of this proof is by contradiction. Assume N is not upper-tailed; there must be variances $v_1 < v_2$ where $v_1 \in N$ and $v_2 \notin N$. The contradiction arises because we assume price is monotonically decreasing in the disclosed variance, so $v_1 < v_2$ implies:

$$P(V \in \hat{N}) \geq P(V = v_1) > P(V = v_2) \geq P(V \in \hat{N}).$$

If price is increasing in the disclosure, then, by an analogous argument, the nondisclosure region is lower-tailed. This completes the proof of Lemma 1.

Proof of Lemma 2

Below, we follow Huang and Litzenberger (1988, Chapter 4.15). Since each investor's budget constraint is binding, we can substitute out the number of bonds through $B_i = W_i^0 - \Sigma_{j=1}^{J} S_{ij}P_j$. Hence:

$$\begin{aligned} W_i &= (W_i^0 - \sum_{j=1}^{J} S_{ij}P_j)\, R_f + \sum_{j=1}^{J} S_{ij}(X_j - 1_{\{v_j \notin N_j\}}C_j) \\ &= W_i^0 R_f + \sum_{j=1}^{J} S_{ij}(X_j - 1_{\{v_j \notin N_j\}}C_j - P_j R_f) \end{aligned}$$

and we can write the investor's problem as:

$$\max_{\{S_{ij}\}_{j=1}^{J}} E[-e^{-a_i(\Sigma_{j=1}^{J} S_{ij}(X_j - 1_{\{v_j \notin N_j\}} C_j - P_j R_f) + W_i^0)}]$$
$$= -e^{-a_i W_i^0 R_f} E[e^{-a_i \Sigma_{j=1}^{J} S_{ij}(\gamma_j F + e_j)}] e^{-a_i \Sigma_{j=1}^{J} S_{ij}(\mu_j - 1\{v_j \notin N_j\} C_j - P_j R_f)}$$

where expectations are taken conditional on investors' beliefs and whether investors know the firm-specific variance for each stock. Evaluating the expectations term yields:

$$E[e^{-a_i \Sigma_{j=1}^{J} S_{ij}(\gamma_j F + e_j)}] = E[e^{-a_i(\Sigma_{j=1}^{J} S_{ij}\gamma_j)F}] \prod_{j=1}^{J} E[e^{-a_i S_{ij} e_j}]$$
$$= e^{(a_i^2/2)(\Sigma_{j=1}^{J} S_{ij}\gamma_j)^2 \sigma^2} \prod_{j=1}^{J} E[e^{[-(a_i S_{ij})^2/2]V_j}].$$

To express each investor's problem in certainty equivalents we change signs, take logs, and divide through by $(-a_i)$:

$$\max_{\{S_{ij}\}_{j=1}^{J}} W_i^0 + \sum_{j=1}^{J} S_{ij}(\mu_j - 1_{\{v_j \notin N_j\}} C_j - P_j R_f)$$
$$- \frac{a_i}{2}\left(\sum_{j=1}^{J} S_{ij}\gamma_j\right)^2 \sigma^2 - a_i^{-1} \sum_{j=1}^{J} \ln(E[e^{[-(a_i S_{ij})^2/2]V_j}]).$$

The first-order condition for an interior maximum for S_{ij} is:

$$\gamma_j\left(\sum_{k=1}^{J} S_{ik}\gamma_k\right)\sigma^2 + S_{ij}H_{ij}(\underline{\hat{N}}, S_{ij}) = a_i^{-1}(\mu_j - 1_{\{v_j \notin N_j\}} C - P_j R_f)$$

where $H_{ij}(\underline{\hat{N}}, S_{ij}) = v_j$ if manager j discloses v_j, and otherwise:

$$H_{ij}(\underline{\hat{N}}, S_{ij}) = \frac{E[V_j e^{[-(a_i S_{ij})^2/2]V_j} | V_j \in \hat{N}_j, \underline{\hat{N}}_{-j}]}{E[e^{[-(a_i S_{ij})^2/2]V_j} | V_j \in \hat{N}_j, \underline{\hat{N}}_{-j}]}$$

where expectations are taken conditional on the disclosure strategies of other managers.

We conjecture, and subsequently verify, that there exists an equilibrium in which $H_{ij}(\underline{\hat{N}}, S_{ij}^*) = H_j(\hat{N}_j)$ when evaluated at the equilibrium demand for stock, S_{ij}^*. We can write all J first-order conditions for investor i's demand for stocks in vector form as:

$$\underline{\underline{M}}\,\underline{S}_i = a_i^{-1}(\underline{m} - \underline{P}R_f)$$

where:

$$\underline{S}_i = (S_{i1}, .., S_{iJ})$$
$$\underline{m} = (\mu_1 - 1_{\{v_1 \notin N_1\}} C, .., \mu_J - 1_{\{v_J \notin N_J\}} C)$$
$$\underline{P} = (P_1, .., P_J)$$

and the elements of the $J \times J$ matrix $\underline{\underline{M}}$ are:

$$M_{jj} = \gamma_j^2\sigma^2 + H_j(\hat{N}_j)$$

and:

$$M_{kj} = \gamma_k\gamma_j\sigma^2 \quad \text{for } k \neq j.$$

Since $H_j(\hat{N}_j) > 0$, the matrix $\underline{\underline{M}}$ is invertible. Hence:

$$\underline{S_i} = \underline{\underline{M}}^{-1}a_i^{-1}(\underline{m} - \underline{P}R_f).$$

Summing over all investors on both sides, we find the aggregate demand for stock j:

$$\sum_{i=1}^{I} \underline{S_i} = \sum_{i=1}^{I} \underline{\underline{M}}^{-1}a_i^{-1}(\underline{m} - \underline{P}R_f).$$

Since each stock is normalized to be in net supply of 1:

$$\underline{1} = \underline{\underline{M}}^{-1}\left(\sum_{i=1}^{I} a_i^{-1}\right)(\underline{m} - \underline{P}R_f)$$

where the left-hand side is a vector of 1s. This reduces to:

$$\underline{\underline{M}}\,\underline{1} = a^{-1}(\underline{m} - \underline{P}R_f)$$

where $a^{-1} = (\Sigma_{i=1}^{I} a_i^{-1})$. Rearranging terms, the equilibrium share prices follow:

$$\underline{P} = [\underline{m} - a\underline{\underline{M}}\,\underline{1}]R_f^{-1}.$$

To verify the conjecture, substitute the equilibrium share prices into the individual investor's demand function:

$$\underline{S_i^*} = \underline{\underline{M}}^{-1}a_i^{-1}(\underline{m} - [\underline{m} - a\underline{\underline{M}}\,\underline{1}]) = \underline{\underline{M}}^{-1}a_i^{-1}a\underline{\underline{M}}\,\underline{1} = a_i^{-1}a\underline{1}.$$

Since $a_i S_{ij}^* = a$ is the same for all investors and each firm j, the conjecture that $H_{ij}(\underline{\hat{N}}, S_{ij}^*) = H_j(\hat{N}_j)$ follows. Defining the scalar as in Gron et al. (2002):

$$\delta_j(\hat{N}_j) \equiv \frac{H_j(\hat{N}_j)}{E[V_j|V_j \in \hat{N}_j]} = \frac{E[V_j e^{(-a^2/2)V_j}|V_j \in \hat{N}_j]}{E[e^{(-a^2/2)V_j}|V_j \in \hat{N}_j]E[V_j|V_j \in \hat{N}_j]} \geq 1$$

completes the proof of Lemma 2.

Proof of Theorem 1

Before proving the theorem, we introduce additional notation to simplify the presentation. Define $D(y)$ as the difference between the price if y is disclosed and the price if y is not disclosed, i.e.:

$$D(y) \equiv P(V|V = y) - P(V|V \geq y) = \frac{\mu - C - ay}{R_f} - \frac{\mu - a\delta(y)E[V|V \geq y]}{R_f}.$$

The argument in $D(y)$ refers to both the disclosed variance as well as investors' inferred disclosure threshold. Clearly with a continuous random variable with a connected support $D(y)$ will be a continuous function. Also, it is clear that the manager wishes to disclose y if $D(y) > 0$. Further, if we define $k = a^2/2$ and $H(y) = \delta(y)E[V|V \geq y]$, then:

$$H(y) = \frac{E[Ve^{kV}|V \geq y]}{E[e^{kV}|V \geq y]} = \frac{\int_y^{\bar{V}} Ve^{kV} f(V|V \geq y)dV}{\int_y^{\bar{V}} e^{kV} f(V|V \geq y)dV} = \frac{\int_y^{\bar{V}} Ve^{kV} f(V)dV}{\int_y^{\bar{V}} e^{kV} f(V)dV}.$$

If H is differentiable, it follows that:

$$H'(y) \equiv \frac{\partial H}{\partial y}(y) = \frac{e^{ky} f(y)}{\left\{\int_y^{\bar{V}} e^{kV} f(V)dV\right\}^2} \left\{\int_y^{\bar{V}} (V - y)e^{kV} f(V)dV\right\} \geq 0.$$

Since $H(y)$ is increasing, there are three cases to consider: $H'(y) > 1$, $0 < H'(y) < 1$, and $H'(y) = 1$.

> **Case 1:** Suppose $H'(y) > 1$ on some interval $[y_L, y_U]$ and assume $H(y_L) - y_L < C/a$ $< H(y_U) - y_U$. Then $D(y_L) < 0 < D(y_U)$. In this case, the partial disclosure determined by threshold y^* is credible, where y^* solves the equation $C/a = H(y^*) - y^*$. Further this equilibrium is unique on the set of partial disclosure strategies if $H'(y) - 1$ is strictly positive, because in that case $D(y) = 0$ holds only at $y = y^*$.
>
> **Case 2:** Suppose $0 < H'(y) < 1$ on some interval $[y_L, y_U]$ and assume disclosure cost solve the condition: $H(y_L) - y_L > C/a > H(y_U) - y_U$. Then $D(y_U) < 0 < D(y_L)$. A in case 1, the partial disclosure strategy determined by threshold y^* is credible, wher y^* solves the equation $C/a = H(y^*) - y^*$. Further, this equilibrium is unique on the set of partial disclosure strategies if $(H'(y) - 1)$ is strictly negative, since then $D(y$ $= 0$ holds only at $y = y^*$.
>
> **Case 3:** Suppose Assumption A5 does not hold, so that $H'(y) = 1$ for all y. In thi case $H(y) - y = H(0)$ for all y. Further full disclosure (or no disclosure) is the unique credible equilibrium as $H(0) > C/a$ (or $H(0) < C/a$). In the case that $H(0) = C/a$, al possible strategies support a credible equilibrium, since the manager is indifferen among all his reporting choices.

This completes the proof of Theorem 1.

Proof of Corollary 1

The result that disclosure results in a higher market price is immediate from Lemma 2, since we have shown that the no-disclosure set is upper-tailed.

Proof of Corollary 2

The result that the share price of the firm is lower under a mandatory disclosure than a partial disclosure regime follows from the fact that the variance term is additively separable in the prices. We add the following additional notation to that used in the proof of theorem 1. Define $ED(y)$ as the difference between the *ex ante* expected share price of the firm under mandatory full disclosure vs. partial disclosure with threshold, y, i.e.:

$$\begin{aligned} ED(y) &\equiv E[P(V)|V \geq 0] - (\Pr\{V|V \leq y\}E[P(V)|V \leq y] \\ &\quad - \Pr\{V|V \geq y\}E[P(V|V \geq y)|V \geq y]) \\ &= \frac{\mu - C - aE[V]}{R_F} - F(y)\left(\frac{\mu - C - aE[V|V \leq y]}{R_F}\right) \\ &\quad - (1 - F(y))\left(\frac{\mu - a\delta(y)E[V|V \geq y]}{R_F}\right) \\ &= (1 - F(y))(a/R_F)\{(\delta(y) - 1)E[V|V \geq y] - C/a\}. \end{aligned}$$

While the difference $(\delta(y) - 1)$ is strictly positive if the threshold, y, is set less than the maximum variance, the fact that a credible partial disclosure equilibrium exists with threshold y ensures that $ED(y) < 0$. Hence, the expected share price under mandatory full disclosure is lower than it is under the credible partial disclosure regime. An analogous argument can be used to show that share price under mandatory no disclosure is higher than it is under the voluntary regime, completing the proof of Corollary 2.

Proof of Theorem 2

Two-fund separation follows by recalling that Lemma 2 established that $S^*_{ij} = a_i^{-1}a$. Define $\bar{\gamma} = \Sigma_{j=1}^J \gamma_j$ and $\bar{\Omega} = \Sigma_{j=1}^J \Omega_j$ where:

$$\Omega_j = \gamma_j\sigma^2\bar{\gamma} + 1_{\{V_j \leq y_j\}}v_j + 1_{\{V_j \geq y_j\}}\delta_j(y_j)E[V_j \geq y_j].$$

The expected excess return on stock j is:

$$E[R_j] - R_f = E\left[\frac{X_j - 1_{\{V_j \leq y_j\}}C_j}{P_j}\right] - R_f = \frac{\{\mu_j - 1_{\{V_j \leq y_j\}}C_j - P_jR_f\}}{P_j} = \frac{a\Omega_j}{P_j}$$

Define the return on the market as $R_m = W_m/W_m^0$ where the economy's terminal wealth is:

$$\begin{aligned} W_m &= \sum_{i=1}^{I} W_i = \sum_{i=1}^{I}\left(W_i^0 - \sum_{j=1}^{J} S^*_{ij}P_j\right)R_f + \sum_{j=1}^{J}\sum_{i=1}^{I} S^*_{ij}(X_j - 1_{\{V_j \leq y_j\}}C_j) \\ &= W_m^0R_f + \sum_{j=1}^{J}(X_j - 1_{\{V_j \leq y_j\}}C_j - P_jR_f) \end{aligned}$$

and the initial wealth of the economy is $W_m^0 = \Sigma_{i=1}^I W_i^0$. The expected excess return on the market is:

$$E[R_m] - R_f = E\left[\frac{W_m - W_m^0 R_f}{W_m^0}\right] = \frac{\sum_{j=1}^{J}(\mu_j - 1_{\{V_j \le y_j\}}C_j - P_j R_f)}{W_m^0} - \frac{\sum_{j=1}^{J} a\Omega_j}{W_m^0} = \frac{a\overline{\Omega}}{W_m^0}$$

To complete the proof of Theorem 2, substitute:

$$\beta_j = \frac{COV[R_j, R_m]}{VAR[R_m]} = \frac{W_m^0}{P_j}\frac{\Omega_j}{\overline{\Omega}}$$

into the CAPM equation in Theorem 2.

> **Lemma 3:** Under Assumption A6, the expected excess return on stock j is increasing and convex in the variance under mandatory disclosure.

Proof of Lemma 3

If manager j discloses, the return on stock j is:

$$R_j(v_j) = \frac{X_j - C_j}{P_j(v_j)} = \frac{\mu_j + \gamma_j F + \varepsilon_j - C_j}{(\mu_j - C_j - a\gamma_j\sigma^2\overline{\gamma} - av_j)}R_f,$$

and the corresponding expected excess return on stock j is:

$$E[R_j(v_j) - R_f] = \frac{a\gamma_j\sigma^2\overline{\gamma} + av_j}{(\mu_j - C_j - a\gamma_j\sigma^2\overline{\gamma} - av_j)}R_f.$$

By differentiation, we find that:

$$\frac{\partial}{\partial v_j}E[R_j(v_j) - R_f] = \frac{a(\mu_j - C_j - a\gamma_j\sigma^2\overline{\gamma} - av_j) + a(a\gamma_j\sigma^2\overline{\gamma} + av_j)}{(\mu_j - C_j - a\gamma_j\sigma^2\overline{\gamma} - av_j)^2}R_f > 0$$

and:

$$\frac{\partial^2}{\partial v_j^2}E[R_j(v_j) - R_f] = \frac{a^2[\mu_j - C_j]R_f}{[P_j(v_j)]^2} < 0.$$

Proof of Corollary 3

$\frac{\partial}{\partial v_j}(E[R_j] - R_f) > 0$ follows from Lemma 3. Note that both $\frac{\partial P_j}{\partial v_j} = -aR_f^{-1}$ and $\frac{\partial \Omega_j}{\partial v_j} = \frac{\partial \overline{\Omega}}{\partial v_j} = 1$ are positive when the manager discloses. Differentiation of the expected excess return on the market:

$$\frac{\partial}{\partial v_j}(E[R_m] - R_f) = \frac{a}{W_m^0}\left(\frac{\partial \Omega}{\partial v_j}\right)$$

and the beta from the proof of Theorem 2 yield the result.

Proof of Corollary 4

It follows immediately from $\frac{\partial P_j}{\partial v_k} = \frac{\partial \Omega_j}{\partial v_k} = 0$ and $\frac{\partial \overline{\Omega}}{\partial v_k} = 1$ that $\frac{\partial}{\partial v_k}(E[R_j] - R_f) = 0$ and $\frac{\partial \beta_j}{\partial v_k} = \frac{W_m^0 \Omega_j}{P_j}(-\overline{\Omega}^{-2})$ when the manager discloses.

Proof of Corollary 5

We show first that the expected excess stock return and the firm's beta are higher under the mandatory full disclosure regime than under voluntary disclosure, then we show how it extends to the comparison between forbidden disclosure and voluntary disclosure. From the proof of Corollary 3, the expected excess return on stock j is:

$$E[R_j(v_j)] - R_f = \frac{aR_f}{P_j(v_j)}(\gamma_j\sigma^2\overline{\gamma} + v_j)$$

if manager j discloses, and:

$$E[R_j(y_j)] - R_f = \frac{aR_f}{P_j(y_j)}(\gamma_j\sigma^2\overline{\gamma} + \delta(y_j)E[V_j|V_j \geq y_j])$$

if manager j does not disclose. As noted we have $P_j(v_j) > P_j(y_j)$ and $v_j < E[V_j|V_j \geq y_j]$, implying that disclosure results in lower excess expected return than nondisclosure in the voluntary regime. Based on Lemma 3, we can apply Jensen's inequality to $R_j(v_j)$ and find that:

$$\begin{aligned} E[R_j(V_j)|V_j \geq y_j] - R_f &= aE\left[\frac{\gamma_j\sigma^2\overline{\gamma} + V_j}{\mu_j - C_j - a\gamma_j\sigma^2\overline{\gamma} - aV_j}\middle| V_j \geq y_j\right]R_f \\ &\geq a\frac{\gamma_j\sigma^2\overline{\gamma} + E[V_j|V_j \geq y_j]}{(\mu_j - C_j - a\gamma_j\sigma^2\overline{\gamma} - aE[V_j|V_j \geq y_j])}R_f \\ &= aE\left[\frac{\gamma_j\sigma^2\overline{\gamma} + V_j}{(\mu_j - C_j - a\gamma_j\sigma^2\overline{\gamma} - aE[V_j|V_j \geq y_j])}\middle| V_j \geq y_j\right]R_f \\ &= E[R_j(E[V_j|V_j \geq y_j])] - R_f \end{aligned}$$

Applying Law of Iterated Expectations:

$$\begin{aligned} &E[R_j(V_j) - R_f|\textit{Full disclose}] \\ &\quad = \Pr\{V_j \leq y_j\}E[R_j(V_j) - R_f|V_j \leq y_j] + \Pr\{V_j \geq y_j\}E[R_j(V_j) - R_f|V_j \geq y_j] \\ &\quad \geq \Pr\{V_j \leq y_j\}E[R_j(V_j) - R_f|V_j \leq y_j] + \Pr\{V_j \geq y_j\}(R_j(E[V_j|V_j \geq y_j]) - R_f) \\ &\quad = E[R_j(V_j) - R_f|\textit{Voluntarily disclose}] \end{aligned}$$

as claimed in the first part of the Corollary. To compare voluntary to no disclosure equilibrium, an analogous proof applies; this completes the proof of Corollary 5.

REFERENCES

Barry, C., and S. Brown. 1985. Differential information and security market equilibrium. *Journal of Financial and Quantitative Analysis* 20 (December): 407–422.

Coles, J. L., U. Loewenstein, and J. Suay. 1995. On equilibrium pricing under parameter uncertainty. *Journal of Financial and Quantitative Analysis* 30 (September): 347–364.

DeMarzo, P. M., and D. Duffie. 1995. Corporate incentives for hedging and hedge accounting. *Review of Financial Studies* 8 (Autumn): 743–771.

Dye, R. A. 1990. Mandatory versus voluntary disclosures: The case of financial and real externalities. *The Accounting Review* 65 (January): 1–24.

Gigler, F. 1994. Self-enforcing voluntary disclosures. *Journal of Accounting Research* 32 (Autumn): 224–240.

Gron, A., B. N. Jorgensen, and N. G. Polson. 2002. Portfolio choice and stochastic volatility. Working paper, Northwestern University, Evanston, IL.

Hodder, L., L. Koonce, and M. L. McAnally. 2001. Behavioral implications of the SEC market risk disclosures. *Accounting Horizons* 15 (March): 49–70.

Huang, C., and R. H. Litzenberger. 1988. *Foundations for Financial Economics*. New York, NY: North-Holland.

Jorion, P. 2002. How informative are Value-at-Risk disclosures? *The Accounting Review* 77 (October): 911–931.

Kanodia, C., A. Mukherji, H. Sapra, and R. Venugopalan. 2000. Hedge disclosures, futures prices and production distortions. *Journal of Accounting Research* 38 (Supplement): 53–82.

Kirschenheiter, M. T., and N. D. Melumad. 2002. Can "big bath" and earnings smoothing coexist as equilibrium financial reporting strategies. *Journal of Accounting Research* 40 (June): 761–796.

Linsmeier, T. J., and N. D. Pearson. 1997. Quantitative disclosures of market risk in the SEC release. *Accounting Horizons* 11 (March): 107–135.

———, D. B. Thornton, M. Venkatachalam, and M. Welker. 2002. The effect of mandated market risk disclosures on trading volume sensitivity to interest rate, exchange rate, and commodity price movements. *The Accounting Review* 77 (April): 343–377.

Melumad, N. D., G. Weyns, and A. Ziv. 1999. Comparing alternative hedge accounting standards: Shareholders' perspective. *Review of Accounting Studies* 4 (December): 265–292.

Penno, M. 1996. Unobservable precision choices in financial reporting. *Journal of Accounting Research* 34 (Spring): 141–149.

Rajgopal, S. 1999. Early evidence on the informativeness of the SEC's market risk disclosures: The case of commodity price risk exposure of oil and gas producers. *The Accounting Review* 74 (July): 251–280.

Stocken, P. 2000. Credibility of voluntary disclosures. *RAND Journal of Economics* 31 (Summer): 359–374.

Verrecchia, R. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5 (December): 179–194.

———. 1990. Information quality and discretionary disclosure. *Journal of Accounting and Economics* 12 (January): 365–380.

Wilson, R. 1968. Theory of syndicates. *Econometrica* 36 (January): 119–131.

THE ACCOUNTING REVIEW
Vol. 78, No. 2
2003
pp. 471–490

The Information Content of the Deferred Tax Valuation Allowance

Krishna R. Kumar
The George Washington University

Gnanakumar Visvanathan
George Mason University

ABSTRACT: An event study demonstrates that disclosures of changes in deferred tax valuation allowances (VA) provide information beyond contemporaneous earnings reports. Prior research shows that, in setting VA, managers consider the extent that taxable income is available from various sources for the realization of deferred tax assets (DTA). Our evidence supports a characterization where investors use VA disclosures to infer management's expectations about DTA, its realizability, and future taxable income available for realization. These findings are more generally relevant for assessing the consequences of reporting standards that require or permit management judgment, especially about future outcomes. In particular, they support the view that discretion can be a vehicle for management to communicate expectations to the benefit of investors.

Keywords: *deferred taxes; valuation allowance; information content; event study; future profits.*

Data Availability: *Data are available from sources identified in the paper.*

I. INTRODUCTION

Deferred tax assets (DTA) represent future tax benefits from deductible temporar differences, tax losses, and tax credits.[1] Realization of DTA is contingent upon th availability of taxable income of appropriate nature in periods in which deductibl

---

[1] Deductible temporary differences typically arise when taxable income (expense) leads (lags) pretax financia income (expense) or when the tax bases of assets increase (liabilities decrease) relative to their balance shee values.

We are grateful to Bill Baber, Linda Bamber, Chris Jones, Fred Lindahl, Jim Patton, Renee Price, and two anonymous reviewers for their comments on the current and previous versions of the manuscript. We also acknowledg helpful comments from workshop participants at American University, University of Cincinnati, George Maso University, The George Washington University, Rutgers University at Camden, Rutgers University at Newark Temple University, the 23rd Annual Congress of the European Accounting Association at Munich, Germany, an the 2002 American Accounting Association Annual Meeting. We also thank Mithu Dey for her research assistance Finally, we gratefully acknowledge the contribution of I/B/E/S International Inc. in providing earnings foreca data, available via the Institutional Brokerage Estimate System. These data have been provided as part of a broa academic program to encourage earnings expectations research.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted March 200
Accepted November 200

temporary differences reverse, or before tax losses or tax credits expire. The availability of such income is not assured, and therefore Statement of Financial Accounting Standards No. 109 (SFAS No. 109, Financial Accounting Standards Board [FASB] 1992) requires a valuation allowance (VA) to reduce DTA when it is "more likely than not" that such assets will expire before they can be realized. Taxable income for the realization of DTA can come from future operations, past operations, or future reversals of taxable temporary differences. Thus, unanticipated disclosures of VA revisions potentially inform investors about DTA, the likelihood of its realization, and about the nature and amount of future taxable income.

For VA change disclosures to be informative, however, managers must use private value-relevant information and expectations when making VA changes. By endorsing the "application of judgment" (rather than providing a rigid set of standards) in determining VA under SFAS No. 109, the FASB apparently both expects and encourages the use of such private information.[2] If managers have little relevant private information, or are unwilling to use private information in setting VA, or if VA disclosures are pre-empted by other information sources, then VA revisions are likely not informative.

Prior research suggests that managers follow SFAS No. 109 in setting VA, and that VA assessments are associated with firm value. In particular, Miller and Skinner (1998) (hereafter MS) find that managers consider the nature of DTA and the extent that DTA are likely to be realized in setting VA. Valuation studies typically indicate negative relations between VA balances and security prices (Amir et al. 1997; Ayers 1998; Amir and Sougiannis 1999). While these studies show how VA is determined and demonstrate value-relevance, they do not investigate whether *disclosures* of VA provide significant and timely *new* information to investors—information that is unavailable from other sources, and can influence investor expectations and move security prices.

The present study addresses this question by using an event-study approach to examine security-price responses to news media disclosures of VA revisions.[3] Such disclosures are typically made concurrently with quarterly earnings. We consider news disclosures—rather than SEC-mandated 10-K disclosures—for three reasons. First, we expect managers to make such disclosures when they judge VA changes to be material. Second, unlike 10-K filings where large and diverse sets of other information are disclosed, VA news disclosures are usually accompanied by quarterly earnings announcements only. Third, except for filings on EDGAR, filing dates for SEC filings are difficult to identify precisely (Easton and Zmijewski 1993).[4] In contrast, news announcements can be dated quite accurately. Thus, compared to SEC filings, news disclosures are relatively uncontaminated sources of VA information and have precisely identified event dates.

Our sample consists of 136 1994 to 1998 news releases. Using standard event-study methodology and OLS regression models, we document negative associations between VA changes and security returns over three-day disclosure event windows. We then consider the nature of information conveyed in the disclosures. In light of the MS findings that managers set VA based on their DTA expectations and the extent that DTA will be realized, we examine whether event-window security returns are associated with revisions in investors' expectations about DTA realizability and future taxable income available to realize DTA. More specifically, the objective is to examine whether VA disclosures cause investors to revise expectations about yet unreported *underlying determinants* of VA. Results support a characterization where investors use the disclosures to infer managers' expectations and revise their assessments of current year DTA and the likelihood of its realization. Furthermore, investors also revise expectations of earnings for the year following the disclosure.[5]

---

[2] SFAS No. 109 provides guidelines for the application of such judgment.

[3] Although we are unaware of other event studies that consider the *informativeness* of VA disclosures, event studies as well as association studies have previously been used to consider the value-relevance of deferred taxes. For example, Beaver and Dukes (1972) use associations with security returns over annual windows to demonstrate the value-relevance of tax deferrals. Givoly and Hayn (1992) demonstrate the value-relevance of deferred tax liabilities by examining security return responses to events leading up to the passage of the Tax Reform Act of 1986.

[4] Effective May 6, 1996, with few exceptions, SEC registrants are required to file on EDGAR. Our study period focuses on the years 1994 through 1998.

The investigation of investor interpretations of VA changes incorporates three features designed to address concerns that the observed security price responses are attributable to other contemporaneous disclosures, rather than to VA changes. First, we include measures of unexpected earnings for the current quarter to control for contemporaneous earnings disclosures. Second, we eliminate observations with disclosures other than VA changes and earnings during a five-day period around the disclosure. Third, we use a control group of quarterly earnings announcements without concurrent VA change disclosures. Results show that investor reassessments of future profit expectations and DTA realizability are more substantial during quarterly earnings announcements with VA disclosures than during other quarterly earnings announcements.

We also consider the possibility that security price effects are attributable to opportunistic use of VA to manage earnings. Tests using proxies for earnings management incentives as additional explanatory variables indicate that our findings are not driven by earnings management.[6]

This study has implications for investors, standard setters, and researchers. In particular, the evidence indicates that VA disclosures—at least those made in news announcements—provide *new* information to investors. Moreover, the evidence suggests that market participants are relatively sophisticated in their interpretation of this information. Results also indicate that allowing discretion in reporting choices, and thereby eliciting management judgment, can be useful. At least in the context of SFAS No. 109, investors appear to understand the judgment process and effectively use disclosures to assess the underlying expectations of managers. However, caution is recommended. First, the evidence is from a single accounting standard, and therefore, it should be considered preliminary. Second, this evidence is based on a sample of firms that make disclosures through the media, prior to 10-K filings. Thus, these results may not generalize to firms that do not use the media to disclose VA changes.

This study also contributes to the literature, which suggests that discretionary accruals can inform investors about future profitability (e.g., Subramanyam 1996). The study shows how a single accrual—the valuation allowance—can inform investors about future profit performance. The ability to communicate future profitability distinguishes VA changes from other accruals, which typically relate to individual financial statement items. Thus, VA disclosures are more directly informative than other accruals, and as such, they are potentially more relevant to investors.

---

[5] Note that by providing direct evidence of investor reassessments of factors underlying VA changes, the present study goes beyond the work of MS (1998). Their study, while identifying factors that determine VA, does not provide evidence on whether investors revise their expectations of these factors based on VA disclosures. Investor reassessments following disclosures occur only to the extent that investors do not fully anticipate the changes.

[6] These findings should not be construed as evidence that VA changes are not used for earnings management in other instances. Firms may be less inclined to use VA changes to manage earnings when they make media disclosures. Thus, our tests are designed as controls for, rather than as direct tests of, earnings management (see Guay et al. [1996] and Phillips et al. [2002] for tests of earnings management). One study that directly focuses on the role of VA changes in earnings management (in the banking industry) is by Schrand and Wong (2001).

The following section discusses provisions of SFAS No. 109 and explanations for the security price effects of valuation allowance disclosures. Section III discusses sample selection. The research design and empirical specifications are presented in Section IV. Section V discusses results of analysis, and Section VI presents the summary.

II. BACKGROUND AND SECURITY PRICE IMPLICATIONS OF VALUATION ALLOWANCE

A feature that distinguishes SFAS No. 109 from prior standards on accounting for deferred taxes is that SFAS No. 109 *requires* the recognition of DTA for all deductible temporary differences, tax loss carryforwards, and tax credits.[7] The underlying rationale is that "the critical (asset) recognition event is the (historical) event that gives rise to deductible temporary differences and carryforwards" (FASB 1992, para. 85). Since realization of these assets is contingent on the availability of taxable income of appropriate character, the standard also requires firms to record a valuation allowance to reduce DTA to the extent that such assets are "more likely than not" to expire unrealized.

Assessments of whether, and in what amount, DTA are realized involve judgments about the amount, timing, and sources of future taxable income. SFAS No. 109 guides such judgments by identifying four income sources that can be used to realize DTA: (1) future reversals of existing temporary differences, (2) future taxable income exclusive of reversing temporary differences, (3) taxable income in prior carryback years if carryback is permitted, and (4) tax-planning strategies.[8] The standard also identifies examples of evidence, both positive and negative, to be considered when estimating future income. Positive evidence, which reduces the need for a VA, includes existing sales contracts and order backlogs, strong historical earnings, and assets with fair values well in excess of book values. Negative evidence, which suggests establishing or increasing the VA, includes current loss carryforwards, a history of expired loss carryforwards, and expected future losses.

The standard encourages managers to use private information to set VA. If managers do so, then VA disclosures can cause investors to revise assessments of the amount and composition of DTA, the extent it is likely to be realized, and about future earnings available for such realization. Security price reactions to VA change disclosures reflect any such revised assessments. On the other hand, managers may lack relevant private information or may be reluctant to use private information if doing so risks higher costs of capital, loss of competitive advantage, or litigation. In such cases, VA disclosures are not informative, and thus, security prices do not respond.

III. SAMPLE SELECTION

We search LEXIS/NEXIS for news releases from 1994 through 1998 that include the keywords "valuation allowance," and report changes in deferred tax VA. When several reports provide identical information for a firm-quarter, we select the first disclosure. This procedure yields 181 VA change disclosures. Of these, we discard eight that do not specify magnitudes and 33 where either necessary quarterly earnings data on the 1999 Standard & Poor's Compustat databases or daily stock returns on the Center for Research on Security

---

[7] Accounting Principles Board Opinion No. 11 (APB No. 11, APB 1967) did not permit recognition of DTA for tax-credit carryforwards and severely limited recognition of tax-loss carryforwards. SFAS No. 96 was even more restrictive, permitting recognition only when realization was assured by the availability of tax-loss carrybacks.

[8] Tax-planning strategies refer to actions taken by managers—for example, investing in taxable instead of tax-exempt securities—to prevent the unrealized expiration of DTA.

Prices (CRSP) databases are not available. We drop four additional observations with potentially confounding news announcements (other than earnings) during the five-day period centered on the news release date. These procedures yield 136 observations where VA changes are disclosed with quarterly earnings announcements (see Table 1, Panel A). One hundred six (106) firms enter the sample once—that is, for only one quarter in the five-year period—and 13 firms appear twice or more. We use this sample, after excluding four outliers, to study aggregate security price reactions to VA disclosures.[9]

Investigations of associations of event-period security price reactions with VA change determinants use a smaller sample because these analyses require additional data, such as changes in DTA and DTL, that are reported annually, but not quarterly, and earnings for eight quarters following the VA disclosures. For nine firms that report VA changes twice in the same fiscal year, analysis using annual data raises concerns that observations are not independent. For these firms, we retain only the earlier of the two disclosures in the sample. Next, we omit three observations that lack required additional data. The remaining 120 observations, representing 113 firms, comprise the sample for the study of the nature of information conveyed in VA disclosures.

The majority of VA changes in the sample (89 of 136) are negative; 69 of 136 disclosures occur in the fourth quarter with the remaining distributed fairly evenly across the other three quarters. Industry distributions, displayed in Table 1, Panel B, follow the classification scheme advanced by Ali and Kumar (1994). As in previous studies (e.g., Amir et al. 1997), manufacturers comprise a relatively large proportion of the sample.

IV. MODEL DEVELOPMENT AND EMPIRICAL SPECIFICATIONS

Are VA Disclosures Informative?

We investigate whether VA disclosures are informative by regressing three-day cumulative market-model residual returns, centered on the disclosure date, on unexpected changes in VA for disclosure quarter t.[10] We control for contemporaneous quarterly earnings releases by including unexpected earnings as an explanatory variable. The empirical specification is:

$$CAR_t = a_0 + a_1\,\Delta VA_t/P_w + a_2\,\Delta E_t/P_w + a_3\,\Delta E_{t-1}/P_w + a_4\,Ret100_t + \varepsilon_{at} \qquad (1)$$

where all variables are as defined in Exhibit 1; a_0–a_4 = estimated parameters; and firm subscripts are suppressed.

In Expression (1), we specify ΔVA_t as the proxy for unexpected disclosure quarter change in VA, based on MS's (1998, 229) finding that valuation allowances typically do not change from year to year. If this finding applies to quarterly VA changes, then expected ΔVA_t is zero.[11]

We include variables ΔE_t and ΔE_{t-1} in Expression (1) so as to reflect unexpected earnings per share (EPS) under Foster's (1977) first-order autoregressive model of quarterly

---

[9] We use Belsley et al.'s (1980) *dffits* statistics to identify outliers.

[10] Results are similar when cumulative residual returns (CAR_t) are computed over five-day windows beginning on day −2, where day 0 is the announcement date.

[11] Since firms are not *required* to report quarterly VA changes, we are unable to document the statistical properties of such changes. However, we investigate annual VA changes by estimating a cross-sectional regression of price-deflated annual per-share VA changes ($\Delta VA_T/P_{T-1}$) on their first lags ($\Delta VA_{T-1}/P_{T-1}$) for the observations in our sample. Estimated intercepts and slopes are not reliably different from zero at conventional levels, thus supporting a simple random walk characterization and the assumption of zero beginning-of-period expectations for annual VA changes.

TABLE 1
Sample Selection and Distribution

Panel A: Sample Selection

News disclosures[a,b] in 1994 through 1998 reported on LEXIS/NEXIS that include the phrase "valuation allowance" and report a deferred tax valuation allowance (VA) change	181
Less news disclosures (events) that do not report the magnitude of the change in VA	(8)
Less events relating to firms that lack necessary earnings and stock returns data on the Compustat and CRSP databases[c]	(33)
Less events excluded because of confounding news disclosures in the five days from day −2 through day +2, where day 0 is the news release date	(4)
Sample for the investigation of security price responses to VA change disclosures[d]	136
Less events excluded from the investigation of the nature of information in VA change disclosures:	
(i) because they are not the first VA change disclosures for the fiscal year;	(9)
(ii) because of the lack of necessary financial data needed to compute explanatory variables; and	(3)
(iii) as outliers identified using Belsley et al. (1980) diagnostics	(4)
Sample for the investigation of the nature of information in VA disclosures[d]	120

Panel B: Industry Distribution

Industry[e]	**Observations Used in Security Price Response Analysis**	**Observations Used in Investigation of Nature of Information**
Basic	8	8
Capital Goods	43	38
Construction	7	6
Consumer Goods	57	49
Energy	3	3
Financial Services	15	14
Transportation	0	0
Utilities	3	2
Total	136	120

---

[a] All news disclosures of VA changes in our sample are made contemporaneously with quarterly earnings announcements.

[b] We include one news disclosure per firm-quarter—the earliest for the quarter.

[c] Earnings data are needed to compute variables that control for the effects of contemporaneous quarterly earnings disclosures.

[d] 47 (40) of 136 (120) disclosures report positive VA changes. 17 (16) disclosures occur in the first fiscal quarter, 17 (16) in the second quarter, 33 (25) in the third quarter, and the remaining 69 (63) in the fourth quarter.

[e] Industry designations are a variation of Sharpe (1982) and follow Ali and Kumar (1994). SIC codes in each industry are:

Basic Industries: 1000–1299, 1400–1499, 2600–2699, 2800–2829, 2870–2899, 3300–3399;
Capital Goods: 3400–3419, 3440–3599, 3620–3629, 3670–3699, 3800–3849, 5080–5089, 5100–5129, 5160–5169, 7300–7399;
Construction: 1500–1999, 2400–2499, 3220–3299, 3430–3439, 5200–5219;
Consumer Goods: 0000–0999, 2000–2399, 2500–2599, 2700–2799, 2830–2869, 3000–3219, 3420–3429, 3600–3619, 3630–3669, 3700–3719, 3850–3899, 3900–3999, 4830–4899, 5000–5079, 5090–5099, 5130–5159, 5180–5199, 5220–5999, 7000–7299, 7400–9999;
Energy: 1300–1399, 2900–2999, 5170–5179;
Finance: 6000–6999;
Transportation: 3720–3799, 4000–4799;
Utilities: 4800–4829, 4900–4999.

earnings changes.[12] Foster's model specifies unexpected quarterly earnings as $\Delta E_t - K\Delta E_{t-1}$, where K is a constant. Thus, a_3 includes K.[13] Following Christie (1987), we scale all financial variables by P_w, the price per share at the beginning of the disclosure window (the end of day -2).

Our variables measure VA change and EPS expectations with error for at least two reasons. First, forecast dates typically precede the disclosure windows by several weeks, and expectations on forecast dates can differ substantially from beginning-of-event-window expectations. Second, estimated forecasts may not accurately reflect unobservable market expectations on forecast dates. Such measurement error can bias estimated coefficients.

We address this measurement error in three ways. First, following Brown et al. (1987), Collins et al. (1994), Lys and Sivaramakrishnan (1988), and others, we include cumulative returns leading up to the event-window ($Ret100_t$) as a proxy for measurement error in expected earnings and VA changes. The objective is to mitigate coefficient bias by including as an explanatory variable a proxy for expectation revisions between the forecast date and the beginning of the disclosure window.[14] Second, we address concerns that Foster's (1977) model does not accurately represent earnings expectations on forecast dates by considering alternative time-series forecasts as well as analysts' forecasts (see footnote 13). Third, we estimate Expression (1) using measurement error proxies starting from alternative forecast dates, specifically 80 and 60 trading days prior to the beginning of the event window and the prior quarter earnings announcement date. We report results only with $Ret100_t$ as they are similar for other measurement error specifications.

Unanticipated VA decreases (increases) signal increases (decreases) in the amounts of DTA likely to be realized and thus provide good (bad) news; therefore, we predict $a_1 < 0$. A substantial quantity of literature on earnings response coefficients predicts a positive

---

[12] We estimate current quarter EPS (E_t) before VA changes as reported EPS minus VA change per share. This assumes that all VA changes flow through earnings into stockholders' equity. However, SFAS No. 109 (para. 36) identifies several transactions for which tax effects (and related VA changes) flow directly into stockholders' equity. If material amounts of VA changes relate to such transactions, then EPS before VA changes are measured with error (see Hanlon and Shevlin [2002] for evidence that such effects can be material in the context of employee stock option-related tax benefits) and coefficient a_2 of ΔE_t can be biased. We are unable to address such measurement error because the details of transactions whose tax effects flow directly into stockholders' equity are typically not disclosed in quarterly earnings announcements or even in quarterly reports.

[13] Results are similar when we use the IMA(1,1) or SRW models to specify unexpected earnings. The IMA(1,1) process provides a parsimonious specification for unexpected earnings in terms of earnings levels and changes, while allowing for transitory unexpected earnings (see Ali and Zarowin 1992). Amir and Lev (1996) and Wahlen (1994) use the levels and changes specification for unexpected quarterly earnings, thus assuming that quarterly earnings follow an IMA(1,1) process in four lags. Results are also similar when we use consensus analysts' forecasts from I/B/E/S as earnings expectation proxies. These forecasts have at least two shortcomings in our context. First, they are not available for well over half of our sample firms. Second, it is unclear whether analysts' earnings forecasts include or exclude VA changes.

[14] We follow Brown et al. (1987) in using $Ret100_t$ as a specification for the measurement-error proxy. The underlying assumption—that beginning-of-quarter time-series forecasts reflect expectations 100 days prior to the event window—appears reasonable considering that (1) a quarter has about 90 calendar days and about 60 to 70 trading days, and (2) expectations reflected in beginning-of-quarter forecasts can be anticipated by security prices even before the prior quarter earnings announcement. Estimating the measurement error proxy over a window of pre-specified (100-day) duration, rather than from actual prior quarter earnings announcement dates to the beginning of the annoumcement window has the added advantage that observations are not lost when the earlier announcement dates are unavailable.

EXHIBIT 1
Variable Definitions

Price and return data are from the CRSP Daily Returns files. Valuation allowance changes (ΔVA_t) and event quarter earnings per share (E_t) are obtained from news disclosures on the WIRES section of LEXIS/NEXIS. Deferred tax assets (DTA), including DTA due to loss carry-forwards and tax credits, and deferred tax liabilities (DTL) are from footnote disclosures in annual report/10-K filings. All other data are from Compustat. Compustat data item numbers are provided in parentheses where applicable.

CAR_t = the three-day cumulative, value-weighted, market-model residual return for days -1 through day $+1$ where day 0 is the VA change disclosure date for quarter t and market-model parameters are estimated over 250 days preceding the event window;
ΔVA_t = the reported per share change in VA for quarter t ($VA_t - VA_{t-1}$);
ΔE_t = ($E_t - E_{t-4}$) where E_t is earnings per share (EPS) before VA change in quarter t and E_{t-i} is EPS for quarter $t-i$ (E_t is Compustat quarterly item # 19);
$\Delta EBT_{(t+1,t+4)}$ [$\Delta EBT_{(t+5,t+8)}$] = the cumulative actual pretax EPS for quarters $t+1$ through $t+4$ [quarters $t+5$ through $t+8$] minus cumulative actual pretax EPS for quarters $t-1$ through $t-4$ [Pretax EPS = Quarterly pretax earnings/split-adjusted common shares outstanding (Compustat quarterly item #23/(item#15 × item#17)];
$\Delta NetDTA_T$ = the per share change in DTA (exclusive of those related to loss carryforwards and tax credits) net of *DTL* for fiscal year T, which includes event quarter t; that is, $NetDTA_T - NetDTA_{T-1}$, where $NetDTA_T$ is DTA minus DTL for year T, and DTA are gross amounts before subtracting VA;
ΔNOL_T = the per share change in DTA due to loss carry-forwards and tax credits for fiscal year T;
$Ret100_t$ = the cumulative raw return over 100 trading days ending on day -2;
$RetNonEvt_t$ = the cumulative raw return over the period from the beginning of the fourth month of fiscal year T to the end of the third month of the following fiscal year but excluding the three-day event window;
$FutRet4_t$ ($FutRet8_t$) = the cumulative raw return for the period from the end of the event window until 45 calendar days after the end of quarter $t+4$ ($t+8$); and
P_w = the closing price per share for day -2.

coefficient a_2 on $\Delta E_t/P_w$.[15] We expect $Ret100_t$ to vary directly with the measurement error in unexpected earnings and inversely with the error in unexpected VA change. If measurement errors are material, we anticipate $a_4 < 0$.[16]

---

[15] Estimating Expression (1) cross-sectionally assumes that all model parameters are constant across sample firms. Such assumptions characterize most cross-sectional security return studies, including Amir and Lev (1996) and Wahlen (1994, 474). An alternative approach that estimates unexpected earnings using forecast parameters from firm-specific earnings time-series—that is, assuming time-invariant forecast model parameters—can serve as a check for robustness. We do not pursue this option since the procedure results in a substantial loss of observations. However, note that findings are robust to the use of analysts' forecasts and across a range of time-series forecast models.

[16] We expect that revisions in earnings (VA change) expectations are associated positively (negatively) with return responses. Thus, correlations between $Ret100_t$ and measurement error in earnings (VA change) expectations are positive (negative). Since these correlation signs coincide with predicted coefficient signs on the respective explanatory variables in Expression (1), the coefficient on $Ret100_t$ in Expression (1) is predicted to be negative (Brown et al. 1987).

What Do Investors Learn from VA Disclosures?

As we note in Section II, VA disclosures potentially communicate value-relevant information about underlying changes in the extent and nature of DTA, its likelihood of being realized and about the availability of future taxable profits for such realization.[17] We investigate evidence of such information by examining associations between event-window stock returns and measures of investor expectation revisions about DTA realizability, including revisions about pretax income for two years following the disclosure quarter. The empirical specification, Expression (2) below, is derived primarily by substituting for ΔVA_t in Expression (1) with proxies for expectation revisions about DTA, the likelihood of its realization, and the availability of future profits for realization (see MS 1998). Firm subscripts are suppressed.

$$\begin{aligned} CAR_t = {} & b_0 + b_1\, \Delta EBT_{(t+1,t+4)}/P_w + b_2\, \Delta EBT_{(t+5,t+8)}/P_w + b_3\, \Delta NetDTA_T/P_w \\ & + b_4\, \Delta NOL_T/P_w + b_5\, \Delta NOL_{T-1}/P_w + b_6\, \Delta E_t/P_w + b_7\, \Delta E_{t-1}/P_w \\ & + b_8\, RetNonEvt_t + b_9\, FutRet4_t + b_{10}\, FutRet8_t + b_{11}\, Ret100_t + \varepsilon_{bt} \end{aligned} \tag{2}$$

where all variables are as defined in Exhibit 1; and b_0—b_{11} = estimated parameters.

The disadvantage of restricting the horizon to two years is that reversals of deductible temporary differences can occur after two years, and earnings beyond two years are relevant for assessing the realizability of the related DTA. On the other hand, earnings for later time periods have a declining influence on VA revisions if reversals or earnings predictability decline as the horizon extends. If so, then the power of statistical tests to detect future earnings' effects declines. Furthermore, sample attrition increases as the horizon is extended. Thus, the two-year horizon is a compromise that trades off a more complete specification with concerns about the consequences of decreasing sample size and statistical power.[18]

Using Foster's (1977) model, we obtain $(\Delta EBT_{(t+1,t+4)} - K'\Delta E_{t-1})$ $[(\Delta EBT_{(t+5,t+8)} - K''\Delta E_{t-1})]$ as the expression for unexpected pretax earnings for quarters $t+1$ through $t+4$ [$t+5$ through $t+8$] relative to the forecast date at the beginning of quarter t.[19] We use these expressions as proxies for event-window expectation revisions for quarter $t+1$ through $t+4$ ($t+5$ through $t+8$) pretax earnings. Thus, we include $\Delta EBT_{(t+1,t+4)}$, $\Delta EBT_{(t+5,t+8)}$ and ΔE_{t-1} (scaled by beginning of event-window price P_w) as independent variables in Expression (2).

Measurement error in using the above expressions as proxies for event-window expectation revisions has two components—(1) expectation revisions from the forecast date in quarter t through the beginning of the event-window, and (2) expectation revisions from the end of the event-window through the disclosure of quarter $t+4$ ($t+8$) pretax EPS. We use raw returns leading up to the event-window (*Ret100*$_t$) to consider the former, and cumulative returns from the end of the event-window until 45 calendar days after the end of quarter t+4 (t+8) following event quarter t (*FutRet4*$_t$ or *FutRet8*$_t$) to proxy for the latter. In computing *FutRet4*$_t$ (*FutRet8*$_t$) we go beyond the end of quarter t+4 (t+8) in order to accommodate that quarter's 10-Q release date.

---

[17] Managers can also make VA changes in response to changes in income from prior periods available to realize loss-carryback-related DTA. Such VA changes will be uninformative since past earnings are "old news." Thus, we do not consider loss-carrybacks in our return specification. We also do not consider income from tax-planning strategies because such actions are limited in scope and their effects are difficult to predict (MS 1998). Visvanathan (1998) finds that very few firms acknowledge the use of such strategies and none in his sample report an income effect.

[18] Findings reported in the following section, that only one-year-ahead earnings are associated with VA changes, support this research design choice. Our specification of future earnings also assumes that managers use forecast horizons that roll forward by one quarter every quarter. An alternative is to assume that managers consider future earnings over *fiscal* year periods. To this end, we also use earnings over the remaining quarters of the event fiscal year and the two following fiscal years as future earnings. Results (not reported) are qualitatively similar to those from Expression (2).

[19] Note that quarter t earnings are not available at the beginning of the announcement window. Hence EPS changes are measured relative to quarters t–1 through t–4.

DTA and DTL are reported annually only; therefore, we use the actual change in DTA net of DTL for fiscal year T (encompassing event quarter t) as the proxy for event-window expectation revisions regarding DTA and the likelihood of its realization through future reversals of taxable temporary differences. Netting DTA and DTL assumes that each dollar of DTL helps realize, and in turn, offsets a dollar of DTA, and that investors are concerned only with the net balance (see MS 1998).

We also follow MS (1998) in distinguishing DTA related to net operating loss credits and carryforwards. The reasoning is that the likelihood of realization of DTA related to loss and credit carryforwards can be lower due to jurisdictional restrictions on income available to realize DTA. Thus, $\Delta NetDTA_T$ proxies for event-window expectation revisions about DTA excluding that from net operating loss carryforwards and credits, and the likelihood of its realization through reversals of taxable temporary differences. ΔNOL_T considers DTA from loss carryforwards and credits.

The use of annual changes ($\Delta NetDTA_T$ and ΔNOL_T) as expectation-revision proxies assumes that $NetDTA_T$ and NOL_T follow random walk processes. Based on cross-sectional regressions, we document significant positive associations for ΔNOL_T (scaled by beginning-of-event-window stock price P_w) with ΔNOL_{T-1}, but no associations between $\Delta NetDTA_T$ and $\Delta NetDTA_{T-1}$. Thus, the data support the random walk assumption for $\Delta NetDTA_T$ but favor a first-order autoregressive process to represent the time-series behavior of ΔNOL_T. As such, we specify the proxy for event-window revisions in investor expectations for the realization of NOL_T as $\Delta NOL_T - k_0 - k_1\,\Delta NOL_{T-1}$, where k_0 and k_1 are parameters of the ΔNOL_T time-series process.[20] We use cumulative raw returns from the end of the third month following the beginning of fiscal year T until the end of the third month following the end of fiscal year T but excluding the three-day event-window, defined as *RetNonEvt*$_T$, as proxies for measurement error in both $\Delta NetDTA_T$ and ΔNOL_T. This assumes that annual reports are released by the end of the third month following the end of the fiscal year.

As in Expression (1), we include current and prior period earnings changes ΔE_t and ΔE_{t-1} to consider event-window expectation revisions for current quarter t earnings, and *Ret100*$_t$ to proxy for measurement error that results from using these measures to capture event-window expectation revisions.

Security price responses to revisions in expectations about future earnings for quarters t+1 through t+4 (t+5 through t+8) following event quarter t are indicated by coefficients b_1 and b_2 in Expression (2). If investors use VA changes to revise expectations about future profits, then, given the evidence that earnings and returns vary directly, we expect $b_1 > 0$ and $b_2 > 0$. Coefficients b_3 and b_4 on $\Delta NetDTA_T/P_w$ and $\Delta NOL_T/P_w$, respectively, reflect responses to expectation revisions regarding DTA and its realization through DTL. Predictions about the signs of these coefficients are predicated on MS's findings that VA changes vary positively with both contemporaneous changes in DTA net of DTL (exclusive of DTA from NOL) and NOL-related DTA.[21] These findings suggest that VA increases (decreases) signal reduced (increased) likelihood of DTA recovery for the year, both for DTA from NOL and from other sources. Such expectation revisions for DTA have security price consequences of opposite sign. Therefore, coefficients b_3 and b_4 on $\Delta NetDTA_T/P_w$ and $\Delta NOL_T/P_w$ are predicted to be negative. We predict coefficient b_6 on $\Delta E_t/P_w$ to be positive. To the extent that variables are measured with error, coefficients b_8 through b_{11} on the measurement-error control variables $RetNonEvt_t$, $FutRet4_t$, $FutRet8_t$ and $Ret100_t$, respectively, are expected to be negative.

---

[20] That is, we assume $\Delta NOL_T = k_0 + k_1\,\Delta NOL_{T-1} + \varepsilon_t$. Thus, unexpected annual ΔNOL_T is specified as $\varepsilon_t = \Delta NOL_T - k_0 - k_1\,\Delta NOL_{T-1}$.

Two features of the research design underlying Expression (2) provide assurance that documented security-return associations with DTA changes and future profitability are attributable to VA disclosures. First, we use a short (three-day) event-window, thus limiting the potential impact of news outside the event-window. Second, we exclude observations with confounding news disclosures other than the quarterly earnings announcement. Even so, in order to address any remaining concerns about spurious effects attributable to unspecified information releases other than VA changes or earnings that might accompany quarterly earnings releases, we conduct additional tests using a control sample of quarterly earnings announcements without VA disclosures. Quarterly earnings announcements by our sample firms that occur in the same fiscal years as the VA disclosures but have no contemporaneous VA disclosures represent a natural control sample because they are drawn from the same set of firm-years as the experimental sample. If revisions of investor expectations of future profits and DTA realization are more substantial for quarters in which VA disclosures occur than for other quarters, then it is unlikely that such revisions are attributable to factors other than the VA disclosures.

In order to compare expectation revisions about DTA/DTL changes and future profitability during quarterly earnings announcements across firm-quarters with and without VA disclosures, we estimate Expression (3) for a pooled set of observations comprising both the experimental (with VA disclosure) and control (without VA disclosure) groups. A dummy variable D_t, coded 1 for observations with VA disclosures, and 0 otherwise, distinguishes the two groups. Other variables in Expression (3) are as in Expression (2).

$$\begin{aligned} CAR_t = {} & c_0 + c_0' D_t + c_1 \Delta EBT_{(t+1,t+4)}/P_w + c_1'(\Delta EBT_{(t+1,t+4)}/P_w) \times D_t \\ & + c_2 \Delta EBT_{(t+5,t+8)}/P_w + c'_2 (\Delta EBT_{(t+5,t+8)}/P_w) \times D_t + c_3 \Delta NetDTA_T/P_w \\ & + c_3' (\Delta NetDTA_T/P_w) \times D_t + c_4 \Delta NOL_T/P_w + c_4' (\Delta NOL_T/P_w) \times D_t \\ & + c_5 \Delta NOL_{T-1}/P_w + c_5' (\Delta NOL_{T-1}/P_w) \times D_t + c_6 \Delta E_t/P_w + c_7 \Delta E_{t-1}/P_w \\ & + c_8 RetNonEvt_t + c_9 FutRet4_t + c_{10} FutRet8_t + c_{11} Ret100_t + \varepsilon_{ct}. \end{aligned} \tag{3}$$

Coefficient sums $(c_1 + c_1')$ $[(c_2 + c_2')]$ reflect market responses to expectation revisions about future earnings in quarters $t+1$ through $t+4$ $[t+5$ through $t+8]$ for experimental sample observations. Similarly, coefficient sums $(c_3 + c_3')$ and $(c_4 + c_4')$ indicate responses to expectation revisions about DTA and its realizability for this sample. Coefficients c_1', c_2', c_3', and c_4' indicate differences in return responses between the experimental and control samples.

---

[21] We verify that associations between annual VA changes and annual changes in DTA (net of DTL and excluding NOL-related DTA), annual changes in NOL-related DTA, and changes in future profitability (relative to the most recent four quarters) are consistent with those documented in MS (1998).

V. RESULTS

Descriptive Statistics

Table 2 provides descriptive statistics. The ratio $|\Delta VA_t/E_t|$ indicates that VA changes have a material impact on reported earnings for the disclosure quarter t. This is not surprising given that firms consider these VA changes large enough to warrant specific disclosure. Disclosure quarter VA changes represent about 70 percent of the VA change for the year (not reported in the table), and are therefore also substantial relative to changes for the remaining quarters. Median (mean) VA changes are about equal (more than twice) in magnitude to quarterly earnings before valuation allowance changes. A majority of VA changes (89 of 136) are negative, indicating expectations that smaller amounts of DTA will go unrealized.

Mean and median event-window cumulative residual returns (CAR_t) are -1.5 percent and -0.8 percent, respectively. Both mean and median price-scaled event-quarter EPS changes ($\Delta E_t/P_w$) relative to quarter $t-4$ are negative. Sample firms appear to be experiencing deteriorating earning performance and this decline perhaps partly explains the negative event-window returns. Mean price-scaled earnings level (E_t/P_w) is also negative, although the median is positive. Almost one-half of sample firms experience losses in the event quarter.[22]

Mean price-scaled change in net DTA other than due to NOL carryforwards ($\Delta NetDTA_T/P_w$) is positive, indicating increasing DTA (net of DTL) on average. Median change is marginally negative, indicating that over one-half of the sample report reductions in $NetDTA_T$ in the event year T. Both mean and median changes in DTA relating to NOL carryforwards ($\Delta NOL_T/P_w$) are positive. These increases in NOL carryforwards are consistent with the relatively poor financial performance of the sample firms. Median levels of future quarter earnings for both quarters $t+1$ through $t+4$ ($EBT_{(t+1,t+4)}/P_w$) and quarters $t+5$ through $t+8$ following the event quarter t ($EBT_{(t+5,t+8)}/P_w$) are positive, but changes relative to the most recent four quarters ($\Delta EBT_{(t+1,t+4)}/P_w$, $\Delta EBT_{(t+5,t+8)}/P_w$) are negative. As such, the majority of sample firms are profitable but experience declining performance after the event quarter. However, mean future earnings levels and changes over both time frames are negative, indicating that a few firms have poor and deteriorating post-event-quarter performance.

Median (mean) market value of equity (MVE) (not reported in Table 2) is \$49 (\$554) million. Although a majority of the sample firms are relatively small, the skewed MVE distribution indicates the presence of large firms.[23] Overall, earnings-announcement disclosers of VA changes are typically smaller than the cross-section of U.S. firms.

Are VA Changes Informative?

Regression results for Expression (1) are in Table 3. Coefficient a_1 on VA changes (-0.309) is reliably negative ($p < 0.01$, one-tailed) in Model 1 (which corresponds to Expression (1)). Thus, increases (decreases) in VA lead to value-relevant downward (upward) revisions in investor expectations regarding the extent that DTA are realized. Coefficients are also reliably less than 1 ($p < 0.01$, one-tailed) (results not reported). If VA changes cause investors to reassess DTA realizability dollar-for-dollar, and if book values of DTA and DTL reflect market values (Guenther and Sansing 2000), then VA change

---

[22] Mean price-scaled event-period EPS (E_t/P_w) falls at the 33rd percentile of the Compustat population for 1994 through 1998. Mean change in price-scaled EPS ($\Delta E_t/P_w$) is at the 31st percentile; so sample firms have both lower and declining performance relative to the median Compustat firm.

[23] Median MVE for our sample is at the 31st percentile of the Compustat population for 1994 through 1998.

TABLE 2
Descriptive Statistics

Variables[a]	N	Mean	Median	25th percentile	75th percentile
Three-day event-window returns					
CAR_t	136	−0.015	−0.008	−0.078	0.034
Measures of valuation allowance changes					
$\Delta VA_t/P_w$	136	0.017	−0.007	−0.021	0.025
$\lvert \Delta VA_t/E_t \rvert$	136	2.250	1.059	0.322	2.306
Measures of contemporaneously disclosed quarterly earnings					
E_t/P_w	136	−0.031	0.004	−0.043	0.017
$\Delta E_t/P_w$	136	−0.096	−0.008	−0.068	0.005
Explanatory variables for security price effects of valuation allowance changes					
$\Delta NetDTA_T/P_w$	120	0.007	−0.001	−0.008	0.010
$\Delta NOL_T/P_w$	120	0.017	0.002	−0.001	0.019
$\Delta EBT_{(t+1,t+4)}/P_w$	120	−0.024	−0.005	−0.077	0.034
$EBT_{(t+1,t+4)}/P_w$	120	−0.041	0.017	−0.099	0.081
$\Delta EBT_{(t+5,t+8)}/P_w$	120	−0.040	−0.011	−0.091	0.051
$EBT_{(t+5,t+8)}/P_w$	120	−0.058	0.005	−0.036	0.069
Other control variables					
$Ret100_t$	136	−0.006	−0.050	−0.245	0.163
$RetNonEvt_t$	120	0.093	−0.033	−0.351	0.438
$FutRet4_t$	120	0.265	−0.032	−0.305	0.298
$FutRet8_t$	120	0.540	−0.158	−0.555	0.786

[a] Variable definitions are in Exhibit 1.

coefficients are 1. Such coefficients can be less than 1, however, (1) if assets and liabilities underlying DTA are recorded at greater than market value (Guenther and Sansing 2000), or (2) if investors discount managers' assessments of VA. They can also be less than 1 to the extent that VA changes reflect reassessments of DTL available for DTA realization, because market values of DTL are, in many instances, discounted relative to book value (Sansing 1998). Coefficients can be biased if there is residual measurement error in the unexpected VA change measure, or if variables correlated with VA changes are omitted.

Coefficient a_2 on $\Delta E_t/P_w$ (0.224), representing event-window return responses to unexpected quarter t earnings, is positive and significant ($p < 0.05$, one-tailed). Thus, consistent with numerous prior studies, contemporaneously reported quarterly earnings are also informative. Finally, coefficient a_4 on $Ret100_t$ is not statistically significant. As such, error in the expectations measures does not appear to be serious.

Tests (not reported) comparing coefficients a_1 on VA changes and a_2 on unexpected earnings indicate no significant differences. Stock prices respond similarly to unexpected changes in earnings due to VA changes and other sources such as current operations. For the analyst, this suggests that earnings due to VA changes are as important as earnings from operations.

Next, we distinguish between VA increases and decreases in Expression (1) by including a dummy variable D_t, coded 1 for VA increases, and 0 otherwise. D_t appears both as a main effect and in interaction with $\Delta VA_t/P_w$ (Model 1A in Table 3). Return responses to

TABLE 3
Regressions of Three-Day Quarterly Earnings Announcement Window Return Residuals on Measures of Unexpected Earnings and Unexpected Valuation Allowance (n = 132[a])

$$\text{Model 1: } CAR_t = a_0 + a_1\,\Delta VA_t/P_w + a_2\,\Delta E_t/P_w + a_3\,\Delta E_{t-1}/P_w + a_4\,Ret100_t + \varepsilon_{at}$$

$$\text{Model 1A: } CAR_t = a'_0 + a''_0\,D_t + a'_1\,\Delta VA_t/P_w + a''_1(\Delta VA_t/P_w) \times D_t + a'_2\,\Delta E_t/P_w + a'_3\,\Delta E_{t-1}/P_w + a'_4\,Ret100_t + \varepsilon_{a't}$$

Variables[b] [Coefficients]	Predicted Coefficient Signs	Model 1 Value	Model 1 t-statistic[c]	Model 1A Value	Model 1A t-statistic[c]
Intercept $[a_0, a'_0]$	?	−0.002	−0.24	−0.022	−1.57
D_t $[a''_0]$	?	–	–	0.029	1.06
$\Delta VA_t/P_w$ $[a_1, a'_1]$	–	−0.309	−3.37***	−0.725	−3.43***
$(\Delta VA_t/P_w) \times D_t$ $[a''_1]$	?	–	–	0.497	1.98**
$\Delta E_t/P_w$ $[a_2, a'_2]$	+	0.224	2.03**	0.299	2.62**
$\Delta E_{t-1}/P_w$ $[a_3, a'_3]$	?	−0.058	−2.28**	−0.049	−1.90*
$Ret100_t$ $[a_4, a'_4]$	–	0.041	1.25	0.048	1.34
Adjusted R^2			0.25		0.27
Coefficient sums				**Value**	**t-statistic[c]**
$\Delta VA_t/P_w$, $(\Delta VA_t/P_w) \times D_t$ $[a_1+a'_1]$ [d]	–			−0.228	−1.82**

***, **, * indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively, for one-tailed tests where coefficient signs are predicted and for two-tailed tests otherwise.

[a] Results are reported after excluding four observations identified as influential using the *dffits* statistic proposed by Belsley et al. (1980).

[b] D_t is coded 1 for observations with positive VA_t^Q, and 0 otherwise. All remaining variables are as defined in Exhibit 1.

[c] All t-statistics are based on White's (1980) heteroscedasticity-consistent standard errors.

[d] The coefficient sum $\Delta VA_t/P_w$, $(\Delta VA_t{'}P_w) \times D_t$ $[a_1+a'_1]$ in Model 1A is the coefficient of ΔVA_t for the 43 (of 132) observations with positive ΔVA_t.

both VA increases (−0.228; $p < 0.05$, one-tailed), indicated by $a_1 + a'_1$, and VA decreases (−0.725; $p < 0.01$, one-tailed), indicated by a_1, are reliably negative. The difference between the two (0.497; $p < 0.05$, two-tailed), indicated by a'_1, is reliably positive and shows that return responses are stronger for VA decreases. The smaller responses to VA increases (which are income-decreasing) are consistent with prior research findings that earnings decreases (bad news) are more timely—that is, anticipated to a greater degree—and less persistent than earnings increases (Basu 1997).

As noted in Table 1, over one-half (69 of 136) of the VA change disclosures are made in the fourth quarter. We investigate differences between return responses to fourth quarter and other quarter VA disclosures by including a dummy variable to distinguish fourth quarter disclosures in a specification similar to Model 1A in Table 3. Results (not reported) indicate no significant differences.[24]

---

[24] Additional tests (not reported), using dummy variables to indicate three of the four disclosure quarters, do not detect significant differences in security price responses based on the quarter of VA change disclosure.

Results are robust to the use of alternative unexpected earnings proxies, winsorizing all independent variables or replacing independent variables with their ranks. In sum, we find that VA change news disclosures provide new value-relevant information comparable in importance—as measured by security price impact—to contemporaneous earnings disclosures.

Tests of the Nature of Information in VA Disclosures

Results for the analysis of factors underlying security price responses to VA change disclosures are in Table 4. Panel A presents results for Expression (2). Model 2 (which corresponds to Expression (2)) is estimated for 120 VA disclosures with available data. Coefficient b_1 on $\Delta EBT_{(t+1,t+4)}/P_w$ (0.093) is positive and significant ($p < 0.05$, one-tailed), but coefficient b_2 on $\Delta EBT_{(t+5,t+8)}/P_w$ (–0.008) is not significant.[25] Thus, investors appear to use VA change disclosures for quarter t to revise their assessments of earnings for quarters $t+1$ through $t+4$, but not for quarters $t+5$ through $t+8$ (the second year). Two explanations come to mind—that managers (on average) do not consider two-year-ahead earnings expectations in setting VA changes, or that investors fully anticipate such expectations. Given the evidence from coefficient b_1 that investors do not fully anticipate one-year ahead earnings, it seems unlikely that they fully anticipate earnings farther out in the future. Thus, it appears more likely that the first explanation, and not the second, holds. Coefficients b_3 (−0.626; $p < 0.01$, one-tailed) and b_4 (−0.482; $p < 0.01$, one-tailed) on $\Delta NetDTA_T/P_w$ and $\Delta NOL_T/P_w$, respectively, are both significantly less than zero. Thus, investors also use VA disclosures to materially revise expectations about DTA for the year and the likelihood of its realization via DTL.

As in Table 3, coefficient b_6 on $\Delta E_t/P_w$ (unexpected current quarter earnings) is positive and significant. The coefficient on $RetNonEvt_t$ is negative and significant ($p < 0.01$, one-tailed) but coefficients on $Ret100_t$, $FutRet4_t$, and $FutRet8_t$ are not. Thus, at least in one instance, the inclusion of a measurement error proxy appears to be useful.

Results for Expression (3) are in Table 4, Panel B. Coefficient c_1' (0.081) on $(\Delta EBT_{(t+1,t+4)}/P_w) \times D_t$ is positive and significant ($p < 0.05$, one-tailed), and indicates that return responses to reassessments of future-profit expectations for quarters $t+1$ through $t+4$ are significantly greater for the VA disclosure sample than for the non-VA disclosure sample. Coefficient c_2' on $(\Delta EBT_{(t+5,t+8)}/P_w) \times D_t$ is not statistically significant, consistent with the absence of evidence in Model 2 (Table 4, Panel A) of expectation revisions relating to quarter $t+5$ through $t+8$ earnings. Coefficients c_3' (−0.427) and c_4' (−0.317) on $(\Delta NetDTA_t/P_w) \times D_t$ and $(\Delta NOL_t/P_w) \times D_t$, respectively, are both negative and significant ($p < 0.05$, one-tailed). As a result, reassessments of DTA and the extent of its realization through DTL are also more substantial for the VA disclosure sample. Coefficient sums $c_1 + c_1'$ (0.106), $c_3 + c_3'$ (−0.615), and $c_4 + c_4'$ (−0.520) estimate return response coefficients on $\Delta EBT_{(t+1,t+4)}/P_w$, $\Delta NetDTA_t/P_w$, and $\Delta NOL_t/P_w$ for VA disclosers and are all statistically significant ($p < 0.01$). These results confirm findings reported in Panel A for the respective variables. Finally, the coefficient on D_t is not significant, indicating that a "VA announcement" effect does not appear to distinguish return responses for the two samples. Overall,

---

[25] Note that future earnings are measured before taxes. Therefore, coefficients of unexpected future quarter earnings (i.e., b_1 or b_2) should be divided by $(1 - T)$, where T is the marginal tax rate, when making any comparisons with coefficients of unexpected current earnings (i.e., b_6).

TABLE 4

Regressions of Three-Day Quarterly Earnings Announcement Window Return Residuals on Measures of Unexpected Earnings and Valuation Allowance Change Determinants

Panel A: Quarterly Earnings Announcements with VA Disclosures (experimental group) (n = 120)

Model 2:
$$CAR_t = b_0 + b_1\,\Delta EBT_{(t+1,t+4)}/P_w + b_2\,\Delta EBT_{(t+5,t+8)}/P_w + b_3\,\Delta NetDTA_T/P_w + b_4\,\Delta NOL_T/P_w + b_5\,\Delta NOL_{T-1}/P_w + b_6\,\Delta E_t/P_w + b_7\,\Delta E_{t-1}/P_w + b_8\,RetNonEvt_t + b_9\,FutRet4_t + b_{10}\,FutRet8_t + b_{11}\,Ret100_t + \varepsilon_{bt}$$

Variables[a] [Coefficients]	**Predicted Coefficient Signs**	**Value**	**t-statistic[b]**
Intercept $[b_0]$	?	0.021	1.87*
$\Delta EBT_{(t+1,t+4)}/P_w$ $[b_1]$	+	0.093	2.04**
$\Delta EBT_{(t+5,t+8)}/P_w$ $[b_2]$	+	−0.008	−0.34
$\Delta NetDTA_T/P_w$ $[b_3]$	−	−0.626	−2.68***
$\Delta NOL_T/P_w$ $[b_4]$	−	−0.482	−2.63***
$\Delta NOL_{T-1}/P_w$ $[b_5]$	?	−0.050	−0.19
$\Delta E_t/P_w$ $[b_6]$	+	0.191	1.50*
$\Delta E_{t-1}/P_w$ $[b_7]$	?	−0.005	−0.06
$RetNonEvt_t$ $[b_8]$	−	−0.074	−3.53***
$FutRet4_t$ $[b_9]$	−	−0.006	−0.98
$FutRet8_t$ $[b_{10}]$	−	−0.003	−0.49
$Ret100_t$ $[b_{11}]$	−	0.059	1.71
Adjusted R^2			0.26

Panel B: Pooled Sample (n = 443[c])

Model 3:
$$CAR_t = c_0 + c_0'\,D_t + c_1\,\Delta EBT_{(t+1,t+4)}/P_w + c_1'(\Delta EBT_{(t+1,t+4)}/P_w) \times D_t + c_2\,\Delta EBT_{(t+5,t+8)}/P_w + c_2'(\Delta EBT_{(t+5,t+8)}/P_w) \times D_t + c_3\,\Delta NetDTA_T/P_w + c_3'\,(\Delta NetDTA_T/P_w) \times D_t + c_4\,\Delta NOL_T/P_w + c_4'\,(\Delta NOL_T/P_w) \times D_t + c_5\,\Delta NOL_{T-1}/P_w + c_5'\,(\Delta NOL_{T-1}/P_w) \times D_t + c_6\,\Delta E_t/P_w + c_7\,\Delta E_{t-1}/P_w + c_8\,RetNonEvt_t + c_9\,FutRet4_t + c_{10}\,FutRet8_t + c_{11}\,Ret100_t + \varepsilon_{ct}$$

Variables[d] [Coefficients]	**Predicted Coefficient Signs**	**Value**	**t-statistic[b]**
Intercept $[c_0]$	?	0.017	2.75***
D_t $[c_0']$	?	−0.003	−0.24
$\Delta EBT_{(t+1,t+4)}/P_w$ $[c_1]$	?	0.025	0.94
$(\Delta EBT_{(t+1,t+4)}/P_w) \times D_t$ $[c_1']$	+	0.081	1.70**
$\Delta EBT_{(t+5,t+8)}/P_w$ $[c_2]$	?	0.035	1.36
$(\Delta EBT_{(t+5,t+8)}/P_w) \times D_t$ $[c_2']$	+	−0.047	−1.33
$\Delta NetDTA_T/P_w$ $[c_3]$	?	−0.188	−1.67*
$(\Delta NetDTA_T/P_w) \times D_t$ $[c_3']$	−	−0.427	−1.98**
$\Delta NOL_T/P_w$ $[c_4]$	?	−0.203	−1.95**
$(\Delta NOL_T/P_w) \times D_t$ $[c_4']$	−	−0.317	−1.81**

(continued on next page)

TABLE 4 (continued)

Variables[d] [Coefficients]	Predicted Coefficient Signs	Value	t-statistic[b]
$\Delta NOL_{T-1}/P_w$ [c_5]	?	0.010	0.07
$(\Delta NOL_{T-1}/P_w) \times D_t$ [c_5']	?	−0.054	−0.18
$\Delta E_t/P_w$ [c_6]	+	0.131	1.91**
$\Delta E_{t-1}/P_w$ [c_7]	?	0.052	1.15
RetNonEvt$_t$ [c_8]	−	−0.037	−3.65***
FutRet4$_t$ [c_9]	−	0.001	0.03
FutRet8$_t$ [c_{10}]	−	−0.001	−0.22
Ret100$_t$ [c_{11}]	−	0.023	1.56
Adjusted R^2			0.13

Coefficient Sums[e]	Predicted Signs	Value	t-statistic[b]
$\Delta EBT_{(t+1,t+4)}/P_w$, $(\Delta EBT_{(t+1,t+4)}/P_w) \times D_t$ [$c_1 + c_1'$]	+	0.106	2.54***
$\Delta EBT_{(t+5,t+8)}/P_w$, $(\Delta EBT_{(t+5,t+8)}/P_w) \times D_t$ [c_2+c_2']	+	−0.012	−0.58
$\Delta NetDTA_T/P_w$, $(\Delta NetDTA_T/P_w) \times D_t$ [c_3+c_3']	−	−0.615	−3.03***
$\Delta NOL_T/P_w$, $(\Delta NOL_T/P_w) \times D_t$ [c_4+c_4']	−	−0.520	−3.27***

***, **, * indicate significance at the 1%, 5% and 10% levels, respectively, for one-tailed tests where coefficient signs are predicted, and for two-tailed tests otherwise.

[a] Variables are as defined in Exhibit 1.

[b] All t-statistics are based on White's (1980) heteroscedasticity-consistent standard errors.

[c] The pooled sample consists of the 120 VA change disclosures with concurrent quarterly earnings announcements in 1994 through 1998 (the experimental group), and 323 quarterly earnings announcements with no concurrent VA change disclosures (the control group). Nondisclosure observations are selected by identifying for each observation in the disclosure sample, the three quarterly earnings announcements in the same fiscal year that do not have a concurrent VA change disclosure. Thirty-seven of the maximum possible 360 nondisclosure observations are lost for lack of data.

[d] D_t is coded 1 for the 120 observations for which a change in valuation allowance is reported concurrently with the quarterly earnings announcement, and 0 otherwise. All other variables are as defined in Exhibit 1. For each variable, the coefficients on interactions with D_t indicate differences between the coefficients on that variable for the experimental and control groups.

[e] $\Delta EBT_{(t+1,t+4)}/P_w$, $(\Delta EBT_{(t+1,t+4)}/P_w) \times D_t$ indicates the sum of the coefficients on $\Delta EBT_{(t+1,t+4)}/P_w$ and $(\Delta EBT_{(t+1,t+4)}/P_w) \times D_t$ [c_1+c_1'], and estimates the coefficient on unexpected cumulative EPS before taxes over quarters $t+1$ through $t+4$ following event quarter t for the experimental sample. The coefficient sum $\Delta EBT_{(t+5,t+8)}/P_w$, $(\Delta EBT_{(t+5,t+8)}/P_w) \times D_t$ [c_2+c_2'] estimates the coefficient on unexpected $t+5$ through $t+8$ cumulative pretax EPS for the experimental sample. The coefficient sums $\Delta NetDTA_T/P_w$, $(\Delta NetDTA_T/P_w) \times D_t$ [c_3+c_3'] and $\Delta NOL_T/P_w$, $(\Delta NOL_T/P_w) \times D_t$ [c_4+c_4'] estimate coefficients on unexpected changes in net deferred tax assets other than due to NOLs and unexpected changes in deferred tax assets due to NOLs, respectively, for the experimental sample.

these results indicate that investor reassessments of DTA, the likelihood of its realization, and the availability of future profits for such realization occur through VA disclosures rather than other unspecified contemporaneous information releases.[26]

---

[26] As an additional test of whether return responses to DTA and future profitability expectation revisions are attributable to VA disclosures, we re-estimate Expression (2) after including $\Delta VA_t/P_w$ as an explanatory variable. We find that coefficients on VA changes ($\Delta VA_t/P_w$) are negative and significant ($p < 0.05$) but those on future profitability and DTA change measures are not (results not reported). Thus, there appears to be little information about DTA and profits available for its realization disclosed during the announcement window that is *incremental* to the VA disclosure. Thus, announcement-window expectation revisions about future profitability and DTA realizability documented in Table 4, Panel A appear to be primarily driven by VA disclosures.

Robustness Tests to Consider Earnings Management Effects as Alternative Explanations

Khalaf (1993) and others suggest that managers can use the discretion available in setting VA to manage reported earnings. Incentives for earnings management arise from efforts to transfer wealth among the firm's stakeholders. If VA changes are used to manage earnings, then security price responses to disclosures of VA changes can be attributable to such wealth transfers and not to the information effects suggested in this paper. We address such concerns (results not reported) by including in Expression (2) firm characteristics that proxy for earnings management incentives. Prior research (see Watts and Zimmerman 1986) shows that firms with higher leverage (higher levels of political visibility) have incentives to report higher (lower) income. Compensation contracts can provide incentives to manage earnings upward (downward) in periods of poor (good) performance. To control for such incentives, we include measures in Expression (2) of firm size, leverage, and a dummy variable *DSMOOTH* that is coded 1 when VA changes reduce the volatility of earnings.[27] All findings remain robust to including these variables.

VI. SUMMARY AND CONCLUSIONS

SFAS No. 109 brought about a significant change in the accounting for income taxes by adopting the asset-liability approach. Two key features of the standard are a more liberal position on the recognition of DTA, and the use of a VA to reduce DTA to the extent that realization of DTA is uncertain. Although the standard provides guidance, the determination of valuation allowances involves considerable managerial judgment. Consequently, the question arises whether and how VA disclosures are useful to investors. Disclosures can be useful if managers use value-relevant, non-public information in setting VA allowances and if such information does not become public prior to the VA disclosures.

This study addresses these questions by examining security price responses to news disclosures of VA changes. The primary result is that, after considering contemporaneous earnings announcements, security price reactions to the news reports are negatively associated with changes in VA. Thus, public disclosures of VA changes are informative.

Further analysis of the nature of information communicated in such disclosures indicates that market participants revise expectations about DTA, its realizability, and about future earnings, for up to four quarters following the event quarter. Security price responses to disclosures suggest that investors understand how managers determine VA under SFAS No. 109 and objectively value deferred tax VA changes.

The findings of the study indicate, quite convincingly, that VA disclosures provide new information to investors. Evidence from prior research with respect to this issue is mixed, and thus inconclusive. By using an event-study approach, the present study offers more powerful tests and thus stronger evidence than previous studies about the informativeness of valuation allowances. More generally, the findings are potentially useful for standard-setting agencies in considering accounting policies and disclosures that rely on the judgment and discretion of management.

---

[27] Increases in VA are income-decreasing and potentially reduce volatility when earnings before VA changes increase. Decreases in VA are income-increasing and reduce volatility when earnings before VA decrease. *DSMOOTH* is coded 1 when VA increases (decreases) coincide with earnings increases (decreases) relative to four prior quarters. Firm size is measured as the natural log of the market value of equity and leverage is measured as the ratio of the book value of total debt to the market value of equity, both at the beginning of year *T*.

REFERENCES

Accounting Principles Board (APB). 1967. *Accounting for Income Taxes*. Accounting Principles Board Opinion No. 11. New York, NY: AICPA.

Ali. A., and P. Zarowin. 1992. The role of earnings levels in annual earnings-returns studies. *Journal of Accounting Research* 30 (Autumn): 286–296.

———, and K. Kumar. 1994. The magnitudes of financial statement effects and accounting choice: The case of the adoption of SFAS No. 87. *Journal of Accounting and Economics* 18 (1): 89–115.

Amir, E., and B. Lev. 1996. Value-relevance of non-financial information: The wireless communications industry. *Journal of Accounting and Economics* 22: 3–30.

———, M. Kirschenheiter, and K. Willard. 1997. The valuation of deferred taxes. *Contemporary Accounting Research* (Winter): 597–622.

———, and T. Sougiannis. 1999. Analysts' interpretation and investors' valuation of tax carryforwards. *Contemporary Accounting Research* (Spring): 1–33.

Ayers, B. 1998. Deferred tax accounting under SFAS No. 109: An empirical investigation of its incremental value-relevance relative to APB No. 11. *The Accounting Review* 73 (April): 195–212.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24: 3–37.

Beaver, W., and R. Dukes. 1972. Interperiod tax allocation, earnings expectations and the behavior of security prices. *The Accounting Review* 47 (April): 320–32.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York, NY: John Wiley & Sons, Inc.

Brown, L., R. Hagerman, P. Griffin, and M. Zmijewski. 1987. An evaluation of proxies for the market's assessment of unexpected earnings. *Journal of Accounting and Economics* 9 (July): 159–193.

Christie, A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting and Economics* 9: 231–258.

Collins, D., S. Kothari, J. Shanken, and R. Sloan. 1994. Lack of timeliness and noise as explanations for the low contemporaneous return-earnings association. *Journal of Accounting and Economics* 18: 289–324.

Easton, P., and M. Zmijewski. 1993. SEC form 10-K/10-Q reports and annual reports to shareholders: Reporting lags and squared market model prediction errors. *Journal of Accounting Research* 31 (Spring): 113–129.

Financial Accounting Standards Board (FASB). 1992. *Accounting for Income Taxes.* Statement of Financial Accounting Standards No. 109. Norwalk, CT: FASB.

Foster, G. 1977. Quarterly accounting data: Time-series properties and predictive-ability results. *The Accounting Review* 52 (1): 1–21.

Givoly, D., and C. Hayn. 1992. The valuation of the deferred tax liability: Evidence from the stock market. *The Accounting Review* 67 (April): 394–410.

Guay, W., S. P. Kothari, and R. Watts. 1996. A market-based evaluation of discretionary accrual models. *Journal of Accounting Research* 34 (Supplement): 83–105.

Guenther, D., and R. Sansing. 2000. Valuation of the firm in the presence of temporary book-tax differences: The role of deferred tax assets and liabilities. *The Accounting Review* 75 (1): 1–12.

Hanlon, M., and T. Shevlin. 2002. Accounting for tax benefits of employee stock options and implications for research. *Accounting Horizons* 16 (1): 1–16.

Khalaf, R. 1993. Read those footnotes! *Forbes* (February 15): 154.

Lys, T., and K. Sivaramakrishnan. 1988. Earnings expectations and capital restructuring: The case of equity-for-debt swaps. *Journal of Accounting Research* 26 (2): 273–299.

Miller, G., and D. Skinner. 1998. Determinants of the valuation allowance for deferred tax assets under SFAS No. 109. *The Accounting Review* 73 (April): 213–233.

Phillips, J., M. Pincus, and S. Rego. 2002. Earnings management: New evidence based on deferred tax expense. Working paper, The University of Iowa, Iowa City, IA.

Sansing, R. 1998. Valuing the deferred tax liability. *Journal of Accounting Research* 36 (Autumn): 357–363.

Schrand, C., and F. Wong. 2001. Evidence on earnings management from the valuation allowance for deferred tax assets under SFAS No. 109. Unpublished manuscript, University of Pennsylvania, Philadelphia, PA.

Sharpe, W. 1982. Factors in New York Stock Exchange security returns, 1931–1979. *Journal of Portfolio Management* 8: 5–19.

Subramanyam, K. R. 1996. The pricing of discretionary accruals. *Journal of Accounting and Economics* 22 (1-3): 250–281.

Visvanathan, G. 1998. Deferred tax valuation allowances and earnings management. *Journal of Financial Statement Analysis* (Summer): 6–15.

Wahlen, J. 1994. The nature of information in commercial loan-loss disclosures. *The Accounting Review* (July): 455–478.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48 (May): 817–838.

THE ACCOUNTING REVIEW
Vol. 78, No. 2
2003
pp. 491–521

Earnings Management: New Evidence Based on Deferred Tax Expense

John Phillips
University of Connecticut

Morton Pincus
Sonja Olhoft Rego
The University of Iowa

ABSTRACT: We assess the usefulness of deferred tax expense in detecting earnings management. Assuming greater discretion under GAAP than under tax rules, and assuming managers exploit such discretion to manage income upward primarily in ways that do not affect current taxable income, then such earnings management will generate book-tax differences that increase deferred tax expense.

Our results provide evidence consistent with deferred tax expense generally being incrementally useful beyond total accruals and abnormal accruals derived from two Jones-type models in detecting earnings management to avoid an earnings decline and to avoid a loss. Only total accruals is incrementally useful in detecting earnings management to meet analysts' earnings forecasts. Deferred tax expense is more accurate than the accrual measures in classifying firm-years as successfully avoiding a loss, whereas no one measure is more accurate in classifying firm-years as avoiding an earnings decline or meeting analysts' forecasts.

Keywords: *earnings management; deferred tax expense; accruals.*

Data Availability: *All data used in this research are from publicly available sources.*

I. INTRODUCTION

In this paper we propose and evaluate the use of deferred income tax expense as a metric for detecting earnings management. Building on the evidence of earnings management in Burgstahler and Dichev (1997) and Mills and Newberry (2001), we investigate the

---

The authors appreciate the comments of two anonymous referees, Ray Ball, Dan Collins, Dan Dhaliwal, Mary Margaret Frank, Michelle Hanlon, Bruce Johnson, Bin Ke, Dawn Matsumoto, Lil Mills, Maria Nondorf, Kathy Petroni, Scott Richardson, Bill Schwartz, Terry Shevlin, D. Shores, workshop participants at the 2002 American Accounting Association Annual Meeting, the Boston Area Research Colloquium, University of Chicago, Columbia University, University of Connecticut, Harvard University, University of Illinois Tax Conference, The University of Iowa, Michigan State University, University of Waterloo, University of Washington, and Washington University, and the programming assistance of Paul Hribar and Hong Xie. Finally, the authors gratefully acknowledge the contribution of Thomson Financial for providing earnings forecast data, available through the Institutional Brokers Estimate System, as part of a broad academic program to encourage earnings expectations research.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted June 2002
Accepted November 2002

usefulness of deferred tax expense in identifying earnings management to meet three earnings targets: (1) to avoid reporting an earnings decline, (2) to avoid reporting a loss, and (3) to avoid failing to meet analysts' earnings forecasts. Detecting earnings management is important in assessing the quality of earnings, and the results of our study should be useful to researchers studying earnings management behavior and to financial analysts in their examination of financial reports. Moreover, evidence that book income is managed in ways that do not affect taxable income contributes to the debate as to whether book income should be the basis for taxation (Yin 2001; Manzon and Plesko 2002).[1]

Prior research has sought to detect earnings management by using various accrual measures as proxies for managerial discretion. However, Guay et al. (1996) demonstrate that accruals derived from five alternative models reflect considerable imprecision, and Bernard and Skinner (1996) argue that abnormal accruals estimated using Jones-type models reflect measurement error due, in part, to the systematic misclassification of normal accruals as abnormal accruals.

We take a different tact and argue that measurement error in accrual metrics used to detect earnings management can be reduced by focusing on deferred tax expense instead of attempting to decompose accruals into normal and abnormal components. Deferred tax expense is a component of a firm's total income tax expense and reflects the tax effects of temporary differences between book income (i.e., income reported to shareholders and other external users) and taxable income (i.e., income reported to the tax authorities) that arise primarily from accruals for revenue and expense items that affect both book and taxable income, but in different periods.

We claim that deferred tax expense can be used to better measure managers' discretionary choices under generally accepted accounting principles (GAAP) because the tax law, in general, allows less discretion in accounting choices relative to the discretion that exists under GAAP (Mills and Newberry 2001; Manzon and Plesko 2002; Hanlon 2002; Joos et al. 2002; Plesko 2002). Hence, we expect that managers seeking to manage earnings to achieve some threshold (e.g., to avoid reporting an earnings decline) do so by exploiting the greater discretion they have for financial reporting purposes *vis-à-vis* tax reporting. Moreover, we assume that managers prefer to manage book income upward without also increasing taxable income. Thus, the exercise of managerial discretion to manage income upward should generate temporary book-tax differences and, hence, deferred tax expense will be useful in detecting such earnings management.[2]

To be sure, firms can manage book income without generating temporary book-tax differences. For example, managers can manage earnings by engaging in a limited set of transactions that create permanent book-tax differences. Managers also can make accrual decisions or take actions that change operating cash flows that affect both book and taxable income simultaneously. However, these latter actions increase current income taxes payable, and if managers take such actions, then we would not detect earnings management using deferred tax expense. Hence, deferred tax expense may not capture all earnings management activity, and it is an empirical question as to whether deferred tax expense is useful for detecting earnings management beyond various accrual measures that have been used in prior research.

---

[1] Basing income taxation on book income rather than taxable income would reduce the level of complexity of income tax rules, and would mean that manipulation of book earnings would induce a tax cost. In addition, accrual-based book income should provide clearer insights into firms' underlying economic activities. Slemrod (2002) argues that income is inherently an accrual-based concept, but that a tax system based on transactions or realizations is easier to administer.

[2] It is possible that firms' tax-planning strategies may lower taxable income without affecting book income. If so, that would also increase deferred tax expense. We assess this possibility in our sensitivity analysis.

We analyze three settings in which the literature argues earnings management likely occurs. The first case we consider is earnings management to avoid an earnings decline. We compare firm-years with zero or slightly positive scaled earnings changes to "just-missed" firm-years (i.e., firm-years with slightly negative earnings changes). The results indicate that increases in deferred tax expense increase the probability of managing earnings to avoid reporting an earnings decline, which is consistent with the argument that deferred tax expense is incrementally useful in detecting earnings management. Total accruals and abnormal accruals estimated using a forward-looking Jones model (Dechow et al. 2002) are also incrementally useful, while abnormal accruals derived from the modified Jones model are not. We find no evidence that any one metric more accurately classifies firm-years as successfully (or unsuccessfully) avoiding an earnings decline. When we examine the impact of firm performance on the results (e.g., Kothari et al. 2002), we find that the accrual measures are no longer significant, while the deferred tax expense results still hold.

The second case we examine is earnings management to avoid a loss, and we compare firm-years with zero or slightly positive scaled earnings levels with a control sample of firm-years with slightly negative earnings. The results suggest that increases in deferred tax expense increase the probability of managing earnings to avoid reporting a loss. Thus, deferred tax expense is also incrementally useful in detecting earnings management in this setting, as are the accrual metrics. However, we find that deferred tax expense is relatively more accurate than each of the accrual measures in classifying firm-years as successfully (or unsuccessfully) avoiding a loss.

Finally, we investigate the usefulness of deferred tax expense in detecting earnings management to avoid failing to meet or beat consensus analysts' earnings forecasts. The literature argues that firms manage earnings upward in this setting, although the evidence in support of this result is mixed (e.g., Schwartz 2001; Matsumoto 2002; Burgstahler and Eames 2002; Dhaliwal et al. 2002). We find no evidence that deferred tax expense or the abnormal accrual metrics detect earnings management to avoid failing to meet or beat analysts' earnings forecasts, whereas total accruals is positively related to the probability that a firm manages earnings in this setting. However, none of the accrual-based metrics or deferred tax expense more accurately classifies firm-years as successfully (or unsuccessfully) avoiding failing to meet or beat analysts' forecasts.

Overall, our results are consistent with the incremental usefulness of deferred tax expense as a metric to detect earnings management. Surprisingly, we find that total accruals is incrementally useful in detecting earnings management activity for each of our three earnings targets, while the abnormal accrual measures do so less consistently. Thus, researchers should consider incorporating deferred tax expense into their research designs to detect the effects of earnings management more fully than relying solely on accrual-based proxies of managerial discretion.

In the next section we develop the intuition underlying our testable hypotheses, summarize relevant prior research, and provide institutional background about the accounting for book-tax differences. Section III describes the empirical design and data. Section IV presents our empirical results, and we conclude in Section V.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

Earnings Management, Discretion, Accruals, and Book-Tax Differences

Earnings management is accomplished through managerial discretion over accounting choices and operating cash flows. Discretion over accruals generally is less observable than management's choices of accounting methods and less costly to implement than altering operating cash flows. Thus, researchers have increasingly used accrual variables to detect earnings management. For example, Healy (1985) uses total accruals to proxy for "discretionary" (i.e., "abnormal") accruals, while Jones (1991) estimates regressions of total accruals on factors reflecting changes in a firm's economic environment to detect earnings management, and uses the residuals to proxy for abnormal accruals.[3,4] Dechow et al. (1995) modify the Jones model to allow for the possibility that managers use discretion to accrue revenues when it is questionable whether revenue recognition criteria have been met.

Dechow et al. (1995) also assess the ability of five accrual models to detect earnings management and find that the modified Jones model is the most powerful in detecting earnings management in a sample of firms the SEC identified for overstating earnings. The evidence in Guay et al. (1996) suggests that only the Jones and modified Jones models produce abnormal accruals that are distinguishable from a random decomposition of earnings and thus consistent with abnormal accruals resulting from managerial decisions to increase and/or smooth income. Moreover, Bernard and Skinner (1996) argue that Jones-type model abnormal accruals systematically misclassify normal accruals as abnormal. Thus, current evidence suggests that accrual variables poorly measure the discretion managers exercise to manage earnings.

We argue that timing (i.e., temporary) book-tax differences will help separate discretion in managers' actions from nondiscretionary choices. As Plesko (2002, 112) notes, "timing differences can arise from different reporting rules under each system, but also because GAAP allows managers greater discretion in determining the amounts of income and expense in each period than does the tax system." For instance, GAAP allows flexibility in estimating the provision for bad debts, while tax rules allow a deduction only for accounts receivable actually written off.[5] Similarly, there is more discretion in choosing useful lives for depreciation under GAAP as compared to the limited flexibility for determining assets' cost recovery periods for tax purposes. There is also more discretion over GAAP revenue recognition. While firms' revenue recognition methods may initially be the same for tax and book purposes, firms that subsequently change to more aggressive methods for financial reporting must continue with their initial tax method unless permission to change is requested and approved by the IRS. There is also discretion over when to recognize unearned revenue as revenue for book purposes, while for tax purposes firms generally must recognize advance payments as income when received. More generally, accruals that require managers' estimates such as post-retirement benefits, restructurings, warranties, and self-insurance reserves generate temporary book-tax differences. In contrast, accruals such as those for accounts receivable, wages payable, and accounts payable that arguably are subject to less managerial discretion typically do not generate temporary book-tax differences.[6]

In addition to having greater discretion for GAAP than for tax purposes, we also assume that firms seeking to manage book income upward prefer to do so without increasing their

---

[3] DeAngelo (1986) uses the change in total accruals, which implicitly assumes that "normal" accruals are constant over time so that a change in accruals reflects abnormal accruals. Dechow (1994) shows that total accruals are mean-reverting; hence, part of the change in total accruals is expected.

[4] DeFond and Jiambalvo (1994) investigate the relation between debt covenant restrictions and accrual choices and report consistent results using both time-series and cross-sectional versions of the Jones model.

[5] GAAP allows for the possibility of increasing income by managing the provision for uncollectible accounts. This results in a smaller deferred tax asset than would otherwise be reported, and thus a larger deferred tax expense.

[6] Firms with different revenue recognition policies for book and tax purposes may have book-tax differences resulting from accounts receivable. Also, accruals for deferred compensation (i.e., compensation paid beyond two and one-half months of year-end) will generate book-tax differences.

tax costs.[7] This assumption, which is analogous to the assumption that it is less costly for managers to manage earnings via accruals rather than operating cash flows, applies both to firms facing nontrivial positive current marginal income tax rates and to firms with a zero current marginal tax rate. Firms in the former group have a current tax incentive to increase book income in ways that do not increase current tax expense, while firms in the latter group that do not have unlimited amounts of loss carryforwards may also seek to minimize the present value of their income taxes.

We thus argue that book-tax differences resulting from accruals that do not increase current taxable income will help separate discretion from nondiscretion. Prior research has linked book-tax differences to earnings management activity. Mills and Newberry (2001) present evidence that firms with earnings management incentives have greater differences between book and taxable income. In particular, public (vs. private) firms, highly leveraged privately held firms, and financially distressed privately held firms all have greater book-tax differences. Based on untabulated results they note that firms reporting slightly positive earnings changes have larger book-tax differences than firms with slightly negative earnings changes. While Mills and Newberry (2001) observe actual book-tax differences using confidential tax return data, we use deferred tax expense, a publicly available measure, as our empirical surrogate for book-tax differences and investigate publicly traded firms' earnings management behavior. We also extend Mills and Newberry (2001) by comparing the abilities of deferred tax expense and accrual-based metrics used in prior research to detect earnings management activity.

Deferred tax expense, our proxy for book-tax differences, is computed in accordance with Statement of Financial Accounting Standards (SFAS) No. 109 (FASB 1992), which takes a balance-sheet approach to accounting for deferred taxes. We thus focus on the period 1994–2000 during which SFAS No. 109 has been in force, while Mills and Newberry (2001) cover the period 1981–1996. SFAS No. 109 defines temporary differences as those differences between the financial accounting and tax bases of assets and liabilities that are expected to reverse in the future, whereas permanent differences will not. Temporary differences can create deferred tax liabilities or deferred tax assets. An increase in deferred tax liabilities is consistent with a firm currently recognizing revenue and/or deferring expense for book purposes relative to its tax reporting, resulting in a future taxable amount. Alternatively, deferred tax assets increase as firms currently recognize expense and/or defer revenue for book *vis-à-vis* tax purposes, thereby producing a future deductible amount. All else equal, firms report higher pretax book income than taxable income when they have increases in their net deferred tax liabilities (defined as the change in deferred tax liabilities less the change in deferred tax assets), and vice versa.[8]

---

[7] Erickson et al. (2002) provide limited evidence on this assumption. They examine a sample of 27 firms cited by the SEC for fraudulent reporting and find evidence that 16 firms increased book income by incurring a tax cost (i.e., through book-tax conforming accruals). For the 16 firms that paid taxes on their earnings overstatement, the tax cost averaged 19 percent of the overstatement.

[8] SFAS No. 109 gives full recognition to deferred tax assets. However, if it is "more likely than not" that a deferred tax asset will not be realized, then a firm must provide a valuation allowance to offset it (FASB 1992, para. 17e). Changes in the valuation allowance account (except for those associated with nonqualified stock option tax deductions, which do not affect earnings) flow through deferred tax expense and affect net income. A firm could decrease its existing valuation allowance account to increase net income and achieve an earnings target, but this would decrease deferred tax expense (or increase a deferred tax benefit). Hence, managing earnings via the valuation allowance would work against finding evidence consistent with deferred tax expense being incrementally useful in detecting earnings management. Prior research finds relatively little evidence that

(continued on next page)

Under SFAS No. 109, the increase (decrease) in net deferred tax liability for a period can equal a firm's deferred tax expense (benefit) for the period, but differences are common. Differences typically occur when firms engage in mergers, acquisitions, and divestitures, or report other comprehensive income items, and can affect deferred tax accounts on the balance sheet without affecting deferred tax expense on the income statement. We focus on deferred tax expense as our empirical surrogate for book-tax differences because it reflects temporary book-tax differences associated with the income statement.[9]

Earnings Thresholds

Burgstahler and Dichev (1997) hypothesize that managers have strong incentives to avoid reporting an earnings decrease and to avoid reporting a loss. They provide evidence of earnings management by documenting a higher frequency of zero or small increases in earnings than expected in cross-sectional distributions of annual scaled earnings changes.[10] They find similar results for zero and slightly positive earnings levels.

Figure 1 shows our replication of Burgstahler and Dichev's (1997) results regarding scaled earnings changes. The unusually high number of observations in the zero and slightly positive earnings change interval and the unusually low frequency of observations in the slightly negative earnings change interval are consistent with their findings. We replicate Burgstahler and Dichev's (1997) scaled earnings levels results in Figure 2. Again, consistent with their findings, there is an unusually high frequency of observations in the zero and slightly positive earnings interval as compared to the slightly negative earnings interval.

We assess the usefulness of deferred tax expense, our empirical proxy for book-tax differences that reflect managerial discretion, to detect earnings management beyond accrual measures used in prior research by investigating whether these variables detect earnings management in the settings Burgstahler and Dichev (1997) consider, i.e., to avoid reporting an earnings decline and to avoid reporting a loss. Hence, we test the following hypotheses:

H1: Deferred tax expense is incrementally useful to accrual measures in detecting earnings management to avoid an earnings decline.

Footnote 8, continued

managers use the valuation allowance account to manage earnings. While Burgstahler et al. (2002) document that managers use the valuation allowance account to report a positive profit, and Schrand and Wong (2002) find evidence that a small sample of banks used the valuation allowance account in 1994 to report a positive earnings change or to meet/beat analysts' forecasts, other papers find no evidence of earnings management via the valuation allowance account (e.g., Bauman et al. 2001; Miller and Skinner 1998).

[9] Deferred tax expense itself conceivably might be decomposed further into normal and abnormal components, but this would require a model of the determinants of deferred tax expense in the absence of earnings management, which we leave for future research. Absent such a model, it might seem that change in deferred tax expense would be a reasonable proxy; however, it does not have a straightforward interpretation. Specifically, $DTE_t - DTE_{t-1} = NDTL_t - 2NDTL_{t-1} + NDTL_{t-2}$, where DTE = deferred tax expense, and $NDTL$ = net deferred tax liability = deferred tax liabilities – deferred tax assets. Under SFAS No. 109, *DTE* is a change variable derived from changes in balance sheet accounts, and is unlikely to follow a random walk. If managers engage in earnings management to increase earnings but not taxable income, then regardless of how the target is defined (e.g., last year's earnings), such earnings management generates book-tax differences that result in a higher *DTE* than would be observed in the absence of such activity. Thus, the level of *DTE*, not the change in *DTE*, is the appropriate variable.

[10] Several papers investigate capital market-based incentives to report earnings increases. Barth et al. (1999) find that firms with sustained periods of zero or positive earnings changes have higher price-earnings ratios than firms unable to sustain earnings growth. Further, the price-earnings multiple increases monotonically for each consecutive year a firm reports nondecreasing earnings but disappears after two years of earnings declines. Other studies (Beatty et al. 2000; Myers and Skinner 2001; Bartov et al. 2002) also support this capital market-based incentive, and consistent with Barth et al. (1999) the incentive to report earnings increases is increasing in a firm's growth opportunities (Skinner and Sloan 2002).

FIGURE 1
Frequency of Firms across Intervals of Scaled Earnings Changes

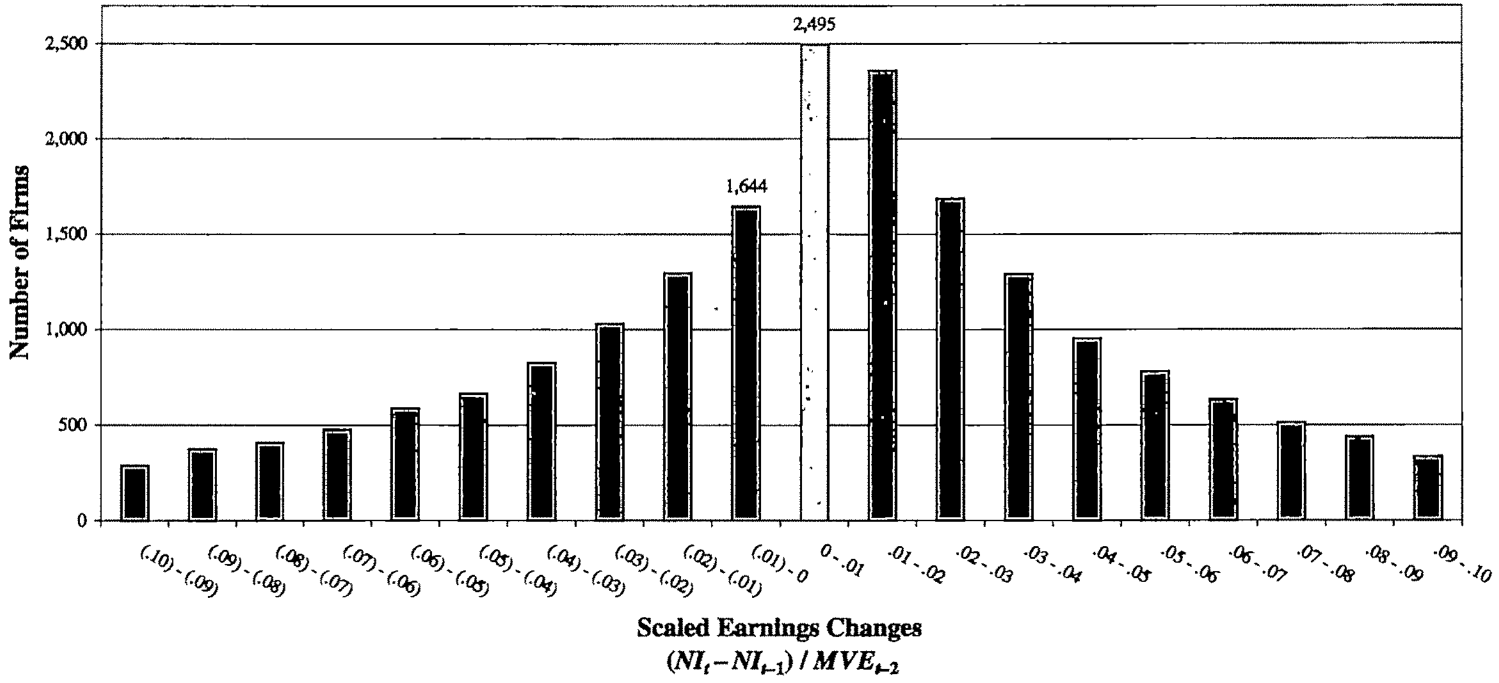
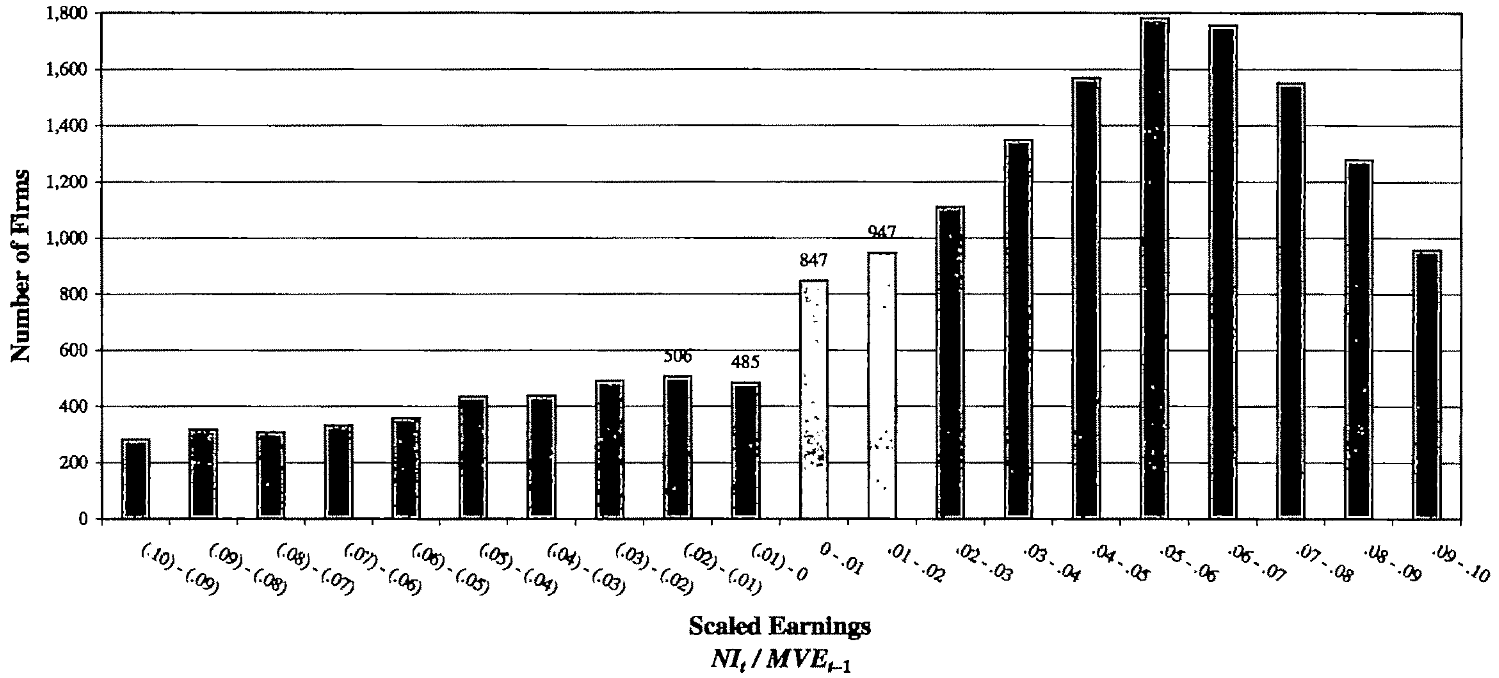


FIGURE 2
Frequency of Firms across Intervals of Scaled Earnings



H2: Deferred tax expense is incrementally useful to accrual measures in detecting earnings management to avoid a loss.

Managers also have incentives to avoid failing to meet or beat analysts' earnings forecasts. For example, Bartov et al. (2002) and Kasznik and McNichols (2002) find that the market rewards firms that meet or beat analysts' forecasts. Figure 3 displays mean analysts' earnings forecast errors by 1 cent per share intervals. Consistent with Burgstahler and Eames (2002), there is a sharply higher frequency of firm-years in the zero and 1 cent per share forecast error intervals as compared to the frequency in the negative 1 cent per share interval. Thus, we consider meeting or beating analysts' forecasts as a third earnings management setting (Degeorge et al. 1999) and hypothesize as follows:

H3: Deferred tax expense is incrementally useful to accrual measures in detecting earnings management to avoid failing to meet or beat analysts' earnings forecasts.

Burgstahler and Eames (2002) find that firms with zero or slightly positive analysts' forecast errors have higher abnormal accruals, computed using the Jones (1991) model. Schwartz (2001) considers the same setting and examines earnings management and earnings guidance given by managers (also see Matsumoto 2002). He reports little evidence of earnings management, but shows that managers guide analysts' earnings forecasts. Unlike avoiding a loss or an earnings decline where the threshold is fixed (e.g., zero or last year's earnings), managers can provide guidance to analysts to induce them to lower their forecasts prior to firms' earnings announcements (e.g., Schwartz 2001; Matsumoto 2002). This complicates our analysis of gauging the usefulness of alternative metrics in detecting earnings management in the analyst forecast setting. However, there is not a consensus in the literature about how to measure managerial guidance (e.g., Schwartz 2001), and we do not control for the impact of guidance. Thus, our examination of earnings management to meet or beat analysts' forecasts is limited and exploratory.[11]

III. EMPIRICAL DESIGN

Research Design

Our primary empirical analyses assess the incremental ability of deferred tax expense and various accrual measures to detect earnings management. We consider three situations in which earnings management likely is present: firm-years with zero or slightly positive earnings changes, firm-years with zero or slightly positive earnings levels, and firm-years where earnings exactly equal or slightly exceed analysts' forecasts.

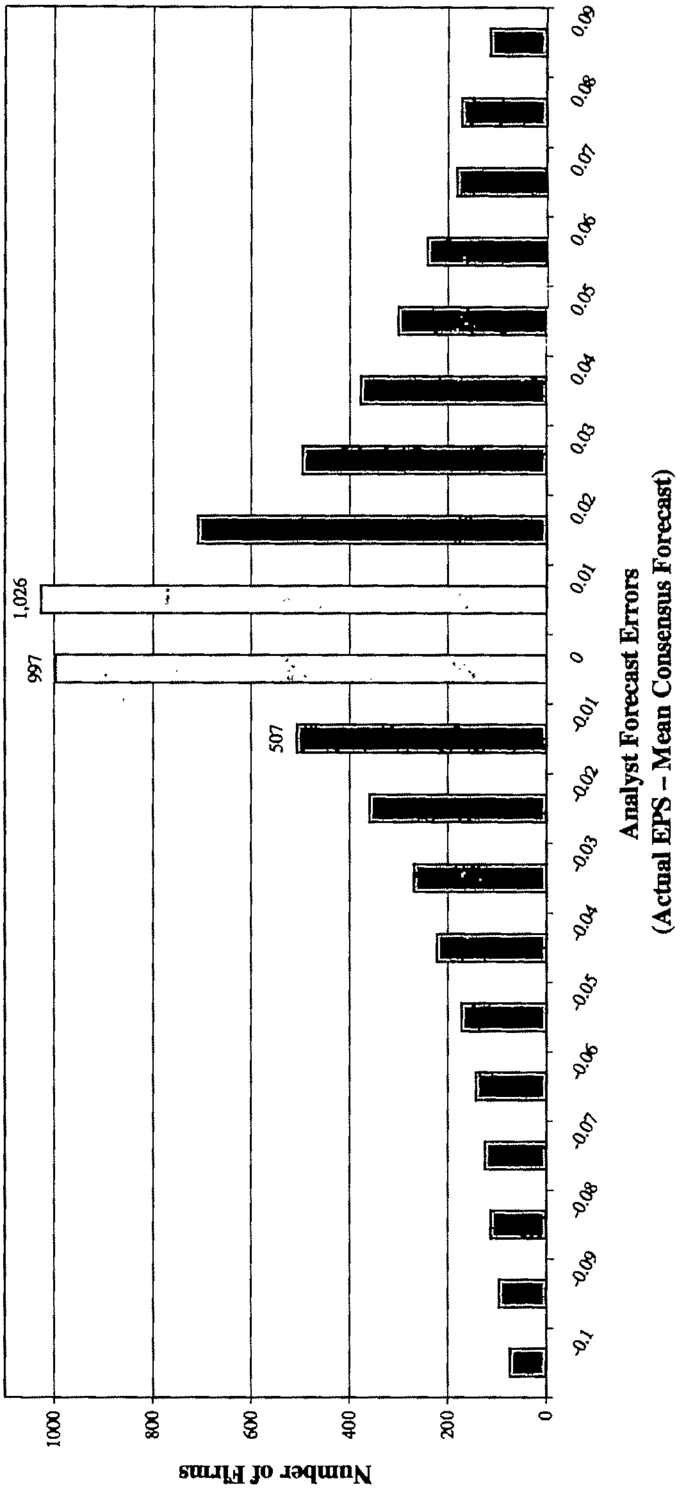
To investigate earnings management to avoid an earnings decline, we estimate the following pooled cross-sectional model using probit regression:

$$EM_{it} = \alpha + \beta_1 DTE_{it} + \beta_2 AC_{it} + \beta_3 \Delta CFO_{it} + \beta_j \Sigma_j Ind_{it} + \varepsilon_{it} \qquad (1)$$

---

[11] In a recent working paper, Dhaliwal et al. (2002) document earnings management to meet analysts' forecasts by examining a specific accrual, namely total tax expense. They hypothesize that the last accrual managers estimate and auditors endorse is tax expense, and this occurs just prior to firms' announcing their annual earnings. Managing total tax expense is, they believe, a last resort for earnings management. In this setting, any guidance of analysts will have already occurred and thus would not confound their analysis. Note, we study pretax earnings management that results in an increase in deferred tax expense, whereas Dhaliwal et al. (2002) study the downward management of income tax expense, which includes the possible manipulation of deferred tax expense, to beat analysts' forecasts.

FIGURE 3
Frequency of Firms across Intervals of Analyst Forecast Errors



where:

EM_{it} = 1 if the change in firm i's net income (annual Compustat data item #172) from year $t-1$ to t divided by the market value of equity at the end of year $t-2$ (annual Compustat data items #25 × #199) is ≥ 0 and <0.01, and 0 if the change in net income ≥ -0.01 and <0;

DTE_{it} = firm i's deferred tax expense (annual Compustat data item #50) in year t, scaled by total assets at the end of year $t-1$;

AC_{it} = a measure of firm i's accruals (see below) in year t;

ΔCFO_{it} = the change in firm i's cash flows from continuing operations (annual Compustat data items #308 − #124) from year $t-1$ to t, scaled by total assets at the end of year $t-1$;

$\Sigma_j Ind_{it}$ = 1 (0) if firm i is (is not) in industry j in year t, based on two-digit SIC codes;

ε_{it} = the error term.

Following Burgstahler and Dichev (1997), EM_{it} equals 1 (0) if firm i reports a scaled earnings change in year t greater than or equal to zero and less than 0.01 (greater than or equal to -0.01 and less than 0) of its beginning-of-year $t-1$ market value of equity.[12] DTE_{it} is the deferred component of firm i's total income tax expense, scaled by its beginning-of-year total assets. We predict the coefficient on DTE_{it} in Equation (1) will be positive and significant, indicating that the probability of earnings management to avoid reporting an earnings decline increases with scaled deferred tax expense.

AC_{it} represents one of three accrual variables (discussed below) used to detect earnings management, and we expect it to have a positive coefficient in the presence of earnings management to avoid an earnings decline. Including both DTE_{it} and AC_{it} in the model allows us to determine the incremental usefulness of each measure in detecting earnings management. In particular, we interpret a positive coefficient on DTE_{it} (AC_{it}) as evidence that it is incrementally useful to AC_{it} (DTE_{it}) in detecting earnings management. We also include ΔCFO_{it} to control for the effect that a change in cash flows from continuing operations has on a firm's status as an earnings management firm. Increases in operating cash flows reflect increases in current performance and reduce the need to manage earnings to achieve a zero or slightly positive earnings change. Finally, we include two-digit SIC industry dummy variables to control for possible differences in the tendency to meet or beat earnings targets across industries.[13]

With regard to the setting of avoiding a loss, consistent with Dechow et al. (2002) we compare firm-years with zero or slightly positive scaled earnings levels to firm-years with slightly negative scaled earnings levels (i.e., a just-missed control sample). We estimate Equation (1), the cross-sectional probit model, after making two changes. First, EM_{it} equals

---

[12] Burgstahler and Dichev (1997) use three scaled earnings change intervals (0–0.005, 0–0.01, 0–0.015) and three scaled earnings levels intervals (0–0.01, 0–0.02, 0–0.03) in their analyses. We use the middle one of their respective three intervals to perform our empirical analyses.

[13] We scale earnings changes by market value of equity to be consistent with Burgstahler and Dichev (1997), but we scale other variables by total assets. In addition to being consistent with prior research that deflates accrual metrics by total assets, we have the concern that deflating by market value would negate the effects we are investigating. That is, since investors bid up the prices of firms that report successive earnings increases and penalize firms that report earnings declines following periods of earnings increases (Barth et al. 1999), then, all else equal, firms that successfully avoid an earnings decline will have higher market values than if they had reported an earnings decline. In turn, this would result in lower values of the test variables for these firms, and the opposite would occur for firms that failed to avoid an earnings decline. Thus, the differences in our scaled test variables for the just-missed and earnings management samples would be minimized, thereby reducing the power of our tests.

1 if firm *i*'s net income (annual Compustat data item #172) in year *t* divided by the market value of equity at the end of year $t-1$ (annual Compustat data items #25 × #199) is at least 0 and less than 0.02, and 0 if greater than or equal to -0.02 and less than 0 (Burgstahler and Dichev 1997). Second, we use the level of cash flows from operations (*CFO*) to control for current performance in the earnings level analysis. We again predict that the coefficient on DTE_{it} will be positive and significant, indicating that the likelihood of managing earnings to avoid reporting a loss increases with deferred tax expense. We interpret a positive coefficient on *DTE* as providing evidence that *DTE* is incrementally useful to the respective accrual-based measure in detecting earnings management in this setting.

Finally, to study the analysts' forecast setting we estimate Equation (1) after redefining EM_{it} to equal 1 if firm *i*'s year *t* analysts' earnings forecast error is zero or 1 cent per share, and 0 if it is -1 cent. We include ΔCFO_{it} as a control for current performance in the analysts' forecast regression, since analysts' forecast errors, like scaled earnings changes in the avoid an earnings decline setting, can proxy for unexpected earnings. A positive and significant coefficient on *DTE* and/or on any of the accrual metrics included in the model would indicate that the likelihood of meeting or beating analysts' forecasts is increasing in *DTE* and/or the accrual metric and would provide evidence of their incremental usefulness in detecting earnings management in this setting.

Accrual Models

We use total accruals (Healy 1985), modified Jones model abnormal accruals (Dechow et al. 1995), and forward-looking abnormal accruals (Dechow et al. 2002) as proxies for accruals that reflect earnings management. Total accruals is earnings from continuing operations minus cash flows from continuing operations:[14]

$$TAcc_{it} = EBEI_{it} - (CFO_{it} - EIDO_{it}) \tag{2}$$

where:

$TAcc_{it}$ = firm *i*'s total accruals in year *t*;
$EBEI_{it}$ = firm *i*'s income before extraordinary items (annual Compustat data item #123) in year *t*;
CFO_{it} = firm *i*'s cash flows from operations (annual Compustat data item #308) in year *t*;
$EIDO_{it}$ = firm *i*'s extraordinary items and discontinued operations from the statement of cash flows (annual Compustat data item #124) in year *t*.

We estimate two different cross-sectional models to derive abnormal accruals. The first is the modified Jones model. Following Dechow et al. (2002), we have:

$$TAcc_{it} = \alpha + \beta_1(\Delta Sales_{it} - \Delta AR_{it}) + \beta_2 PPE_{it} + \xi_{it} \tag{3}$$

where:

---

[14] Following Hribar and Collins (2002), we use data from the cash flow statement rather than from successive balance sheets to estimate accrual measures.

$\Delta Sales_{it}$ = the change in firm i's sales (annual Compustat data item #12) from year $t-1$ to t;

ΔAR_{it} = the change in firm i's accounts receivable from operating activities from year $t-1$ to t (annual Compustat data item #302);

PPE_{it} = firm i's year t gross property, plant, and equipment (annual Compustat data item #7);

ξ_{it} = the error term.

We scale all variables by beginning-of-year total assets (annual Compustat data item #6). Subtracting ΔAR_{it} modifies the Jones (1991) model so that credit sales are assumed to be discretionary. Following Dechow et al. (1995), we estimate the modified Jones model using only the subsample of firm-years in which we assume there is no earnings management (i.e., non-*EM* = 1 firm-years) and, accordingly, we exclude ΔAR_{it} from the estimation of Equation (3). We do this separately for each two-digit SIC group-year with at least ten firms. We then use the estimated parameters in Equation (3) to compute abnormal accruals (denoted *AbAccMJ*).

We also estimate abnormal accruals using Dechow et al.'s (2002) forward-looking model:

$$TAcc_{it} = \alpha + \beta_1(\Delta Sales_{it} - (1 - k)\Delta AR_{it}) + \beta_2 PPE_{it} + \beta_3 TAcc_{it-1} + \beta_4 GR\_Sales_{t+1} + \xi_{it} \quad (4)$$

where:

k = the slope coefficient from a regression of ΔAR_{it} on $\Delta Sales_{it}$;

$TAcc_{it-1}$ = firm i's total accruals from the prior year, scaled by year $t-2$ total assets;

GR_Sales_{it+1} = the change in firm i's sales from year t to $t+1$, scaled by year t sales;

ξ_{it} = the error term.

The forward-looking model includes three adjustments to the modified Jones model. First, rather than assuming all credit sales are discretionary, the model treats part of the increase in credit sales as expected (a normal accrual) by regressing ΔAR_{it} on $\Delta Sales_{it}$ and winsorizing the estimated parameter k so it ranges from 0 to 1. Hence, the change in sales in Equation (4) is reduced by less than 100 percent of the increase in receivables. Second, a portion of total accruals is assumed to be predictable and captured by including last year's accruals (i.e., lagged total accruals) in the model. Third, the modified Jones model treats increases in inventory made in anticipation of higher sales as an abnormal accrual reflecting earnings manipulation rather than as a rational increase in inventory (e.g., Hunt et al. 1996). Including future sales growth corrects for such misclassifications, although it means the forward-looking model uses future period data to estimate current period normal and abnormal accruals. Under the assumption of no earnings management, we estimate the model by excluding $(1 - k)\Delta AR_{it}$ and using non-*EM* = 1 firm-years for each two-digit SIC group-year. We use the resulting parameter estimates in Equation (4) to compute abnormal accruals (denoted *AbAccFL*).

Samples

SFAS No. 109 became effective in 1993 and substantially altered GAAP for income tax reporting. To assure consistent financial reporting, we only include firm-years for the period 1994–2000.[15] We require sample firms to be incorporated in the U.S. because foreign firms face different financial accounting standards, tax rules, and incentives than U.S. firms. We exclude utilities (SIC codes 4900–4999) and financial institutions (SIC codes 6000–6099) because regulated firms may have different incentives regarding earnings management than other businesses,[16] and we exclude mutual funds (SIC code 6726), trusts (SIC code 6792), REITs (SIC code 6798), limited partnerships (SIC code 6799), and other flow-through entities (SIC code 6795) since these firms do not account for income tax expense. A firm-year must have non-missing data for the variables needed in the analysis. To control for extreme observations, we delete firm-years having deferred tax expense below the 1st percentile or above the 99th percentile. Consistent with DeFond and Subramanyam (1998) and Dechow et al. (2002), we also delete firm-years having a scaled accrual measure greater than 100 percent (in absolute value) of total assets.

We compare firm-years with zero or slightly positive scaled earnings changes to firm-years with slightly negative scaled earnings changes. Our selection procedures generate samples that range from 3,352 to 4,139 firm-years in the probit analysis, depending on the accrual variable to which *DTE* is being compared. Approximately 60 percent of the firm-years have scaled earnings changes that are zero or slightly positive (i.e., greater than or equal to zero and less than 0.01 of the market value of equity), and they comprise our earnings management (i.e., $EM = 1$) sample for this analysis. Firm-years with scaled earnings changes that are greater than or equal to -0.01 and less than zero of the market value of equity comprise the just-missed (i.e., $EM = 0$) control sample.

We employ the same criteria to select the earnings management and just-missed samples with regard to avoiding a loss. Our samples range in size from 2,254 to 2,785 firm-years for the probit analysis, depending on the accrual model being considered, with approximately 64 percent of the firm-years having scaled earnings that are zero or slightly positive.

For our analysis of analysts' earnings forecasts, we obtain forecast and actual earnings data from Thomson Financial (I/B/E/S), and use the last mean forecast prior to annual earnings announcements over the 1994–2000 period. We compute analysts' forecast errors (AFEs) as actual earnings minus analysts' mean earnings forecasts per share, and keep firm-years with AFEs between ±1 cent per share, inclusive. We then merge the AFEs with our database of firm-year data needed to do the empirical analyses (e.g., *DTE* and data to compute accrual metrics). Finally, we delete firms having I/B/E/S stock split adjustment factors more than three to eliminate likely *EM* misclassifications due to rounding problems (see Baber and Kang 2002; Payne and Thomas 2002).[17] This yields samples of 2,179 and 2,530 observations, depending on the accrual metric computed. Approximately 80 percent of the firm-year observations are in the $EM = 1$ (i.e., the meet or just beat analysts' earnings forecasts) group.[18]

IV. RESULTS

Graphical Evidence: Deferred Tax Expense and Earnings Thresholds

For descriptive purposes, we provide evidence of the relation between scaled deferred tax expense and, respectively, earnings changes, earnings levels, and AFEs. Note, however, our primary analyses focus on the intervals around the earnings thresholds we are considering.

---

[15] 1994 is the first year we can compute change variables. We lose observations from 2000 when estimating the forward-looking model since we need to use one-year-ahead sales.

[16] Also, Compustat does not report deferred tax accounts for financial institutions.

[17] Our results are not sensitive to (1) including firms with stock split adjustment factors in excess of 3, which tends to include more large firms, or (2) expanding the range of AFEs from ±1 cent to ±5 cents per share.

[18] Data in Schwartz (2001, Figure 1) indicate an increasing proportion of firms in the meet or just beat group during 1994–2000, ranging between approximately 65 percent and 88 percent.

Figure 4 presents a histogram of mean *DTE* by scaled earnings change intervals that have a width of 0.01 of market value and range from −0.10 to 0.10. Note that the mean *DTE* for the −0.01 to less than 0 interval (i.e., the just-missed interval) is 0.0005, whereas the mean *DTE* is 0.0015 in the zero and slightly positive earnings change interval. Consistent with H1, mean *DTE* is higher for firm-years that just avoid an earnings decline. We also graphed each of the accrual metrics by scaled earnings change intervals (not shown). Mean *TAcc* increases slightly in the zero or slightly positive earnings change interval *vis-à-vis* the just-missed interval, while the corresponding means for both abnormal accrual measures are flat around zero scaled earnings changes.

We display the histogram of mean *DTE* by scaled earnings levels in Figure 5. (To facilitate comparison with the other figures, scaled earnings intervals in Figure 5 have a width of 0.01 of market value and range from −0.10 to 0.10.) Likely due to the tax benefits of losses, the mean *DTE* is negative for all loss intervals, but becomes less negative when earnings are zero or slightly positive. More specifically, mean *DTE* is −0.0047 in the slightly negative earnings interval and −0.0034 in the adjoining zero and slightly positive earnings interval, which is consistent with H2. Graphs for accruals (not shown) suggest that mean *TAcc* also becomes less negative while the means of both abnormal accruals metrics sharply increase.

Figure 6 shows the histogram of mean *DTE* by AFE intervals. Mean *DTE* is 0.0002 in the −1 cent per share interval, and 0.0005 and 0.0008, respectively, in the zero and 1 cent per share intervals. This difference is consistent with H3. The graph of total accruals (not shown) also shows higher mean values in the meet or beat intervals relative to the just-missed interval, while the graphs of the abnormal accrual measures do not.

Descriptive Statistics and Univariate Analysis

Panel A of Table 1 presents summary statistics for our comparison of firm-years with zero or slightly positive earnings changes vs. firm-years with slightly negative earnings changes. For the *EM* = 1 sample, the mean *DTE* is 0.0015, or 0.15 percent of beginning-of-year total assets (median = 0.0001), with values ranging from −9.33 percent to 7 percent of total assets. The mean *TAcc* is −0.0443 or −4.43 percent of beginning-of-year total assets (median = −0.0393), and the range is from −76.42 percent to 89.90 percent. Because it is a pretax amount, total accruals is substantially larger in magnitude then *DTE*, and it is negative because it includes depreciation and amortization accruals. In the just-missed control sample, mean *DTE* is 0.0005 (median = 0.0000) and mean *TAcc* is −0.05 (median = −0.0446). Both abnormal accruals variables have negative means and medians in both samples.

We statistically compare the two samples on a univariate basis (p-values are two-tailed). We expect that if firms manage earnings upward to avoid reporting an earnings decline, then earnings management metrics should reflect this activity. In particular, we expect greater deferred tax expense and greater accrual values in earnings management firm-years than in control firm-years. The results indicate that the mean and median for both *DTE* and *TAcc* are significantly larger in the *EM* = 1 sample of firm-years that just avoid an earnings decline than in the just-missed control sample. However, we do not observe significantly larger abnormal accruals for the *EM* = 1 firm-years. Panel A also indicates that changes in operating cash flows are reliably larger for *EM* = 1 firm-years, supporting its inclusion as a control variable in Equation (1).

FIGURE 4

Distribution of Deferred Tax Expense across Intervals of Scaled Earnings Changes

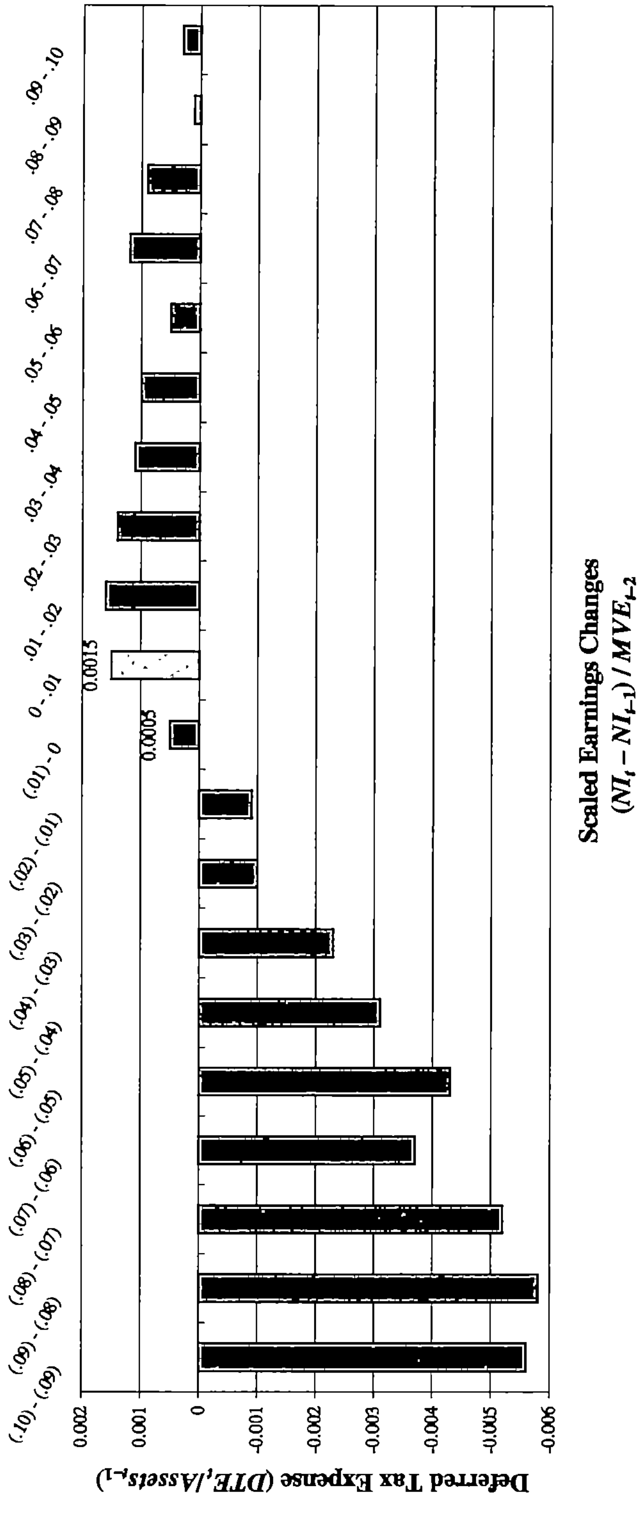


FIGURE 5
Distribution of Deferred Tax Expense across Intervals of Scaled Earnings

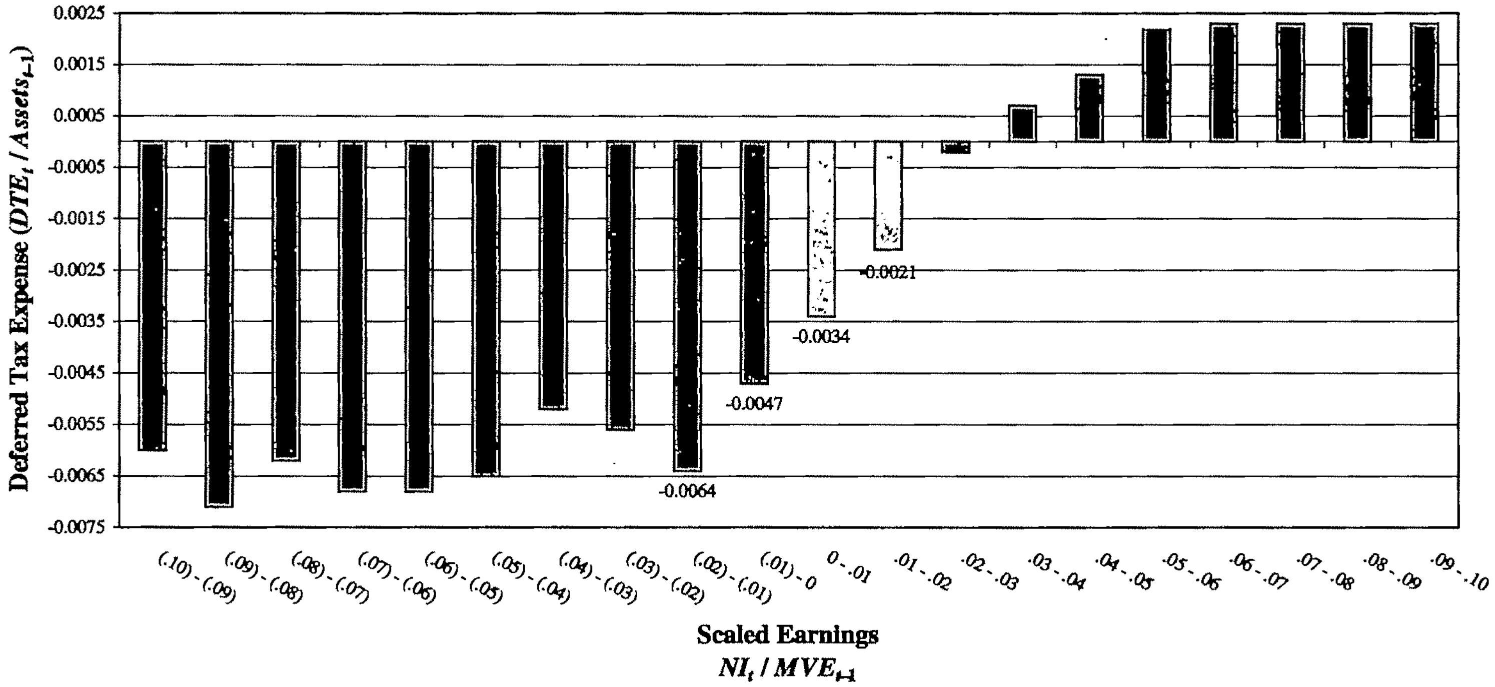


FIGURE 6
Distribution of Deferred Tax Expense across Intervals of Analyst Forecast Errors

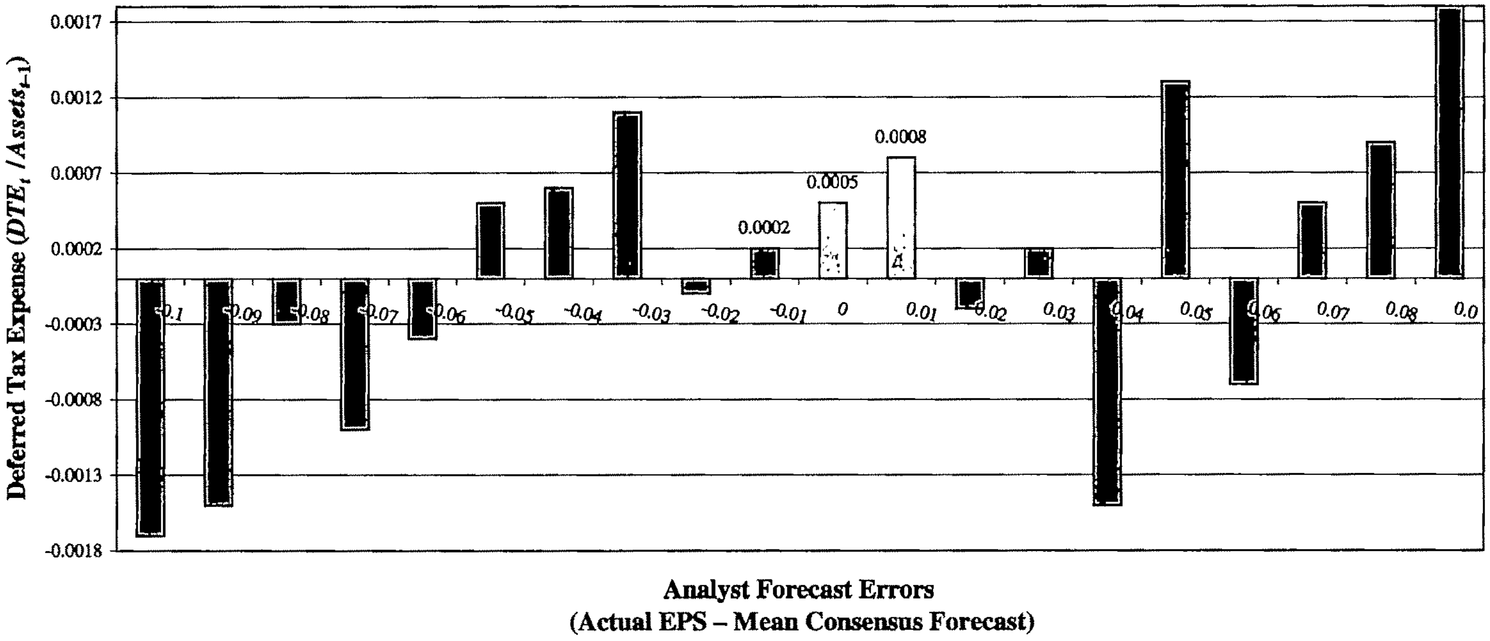


TABLE 1
Descriptive Statistics and Univariate Analysis

Panel A: Earnings Management to Avoid an Earnings Decline Samples: Zero and Slightly Positive Earnings Changes ($EM_1 = 1$) vs. Slightly Negative Earnings Changes ($EM_1 = 0$), where $EM_1 = 1$ Firm-Years Have Scaled Earnings Changes [$(NI_t - NI_{t-1})/MVE_{t-1}$] of at Least 0 and Less Than 0.01 and $EM_1 = 0$ Firm-Years Have Scaled Earnings Changes of at Least −0.01 and Less Than 0

					Percentiles			
EM_1 = 1 Firm-Years	**n**	**Mean**	**Std. Deviation**	**Maximum**	**75th**	**50th**	**25th**	**Minimum**
DTE	2,495	0.0015*	0.0128	0.0700	0.0066	0.0001*	−0.0027	−0.0993
TAcc	2,495	−0.0443*	0.0979	0.8990	−0.0034	−0.0393*	−0.0815	−0.7642
AbAccMJ	2,495	−0.0126	0.1606	0.9888	0.0505	−0.0109	−0.0717	−0.9969
AbAccFL	2,057	−0.0009	0.1368	0.9431	0.0533	−0.0074	−0.0619	−0.9493
ΔCFO	2,495	0.0215*	0.1233	1.3465	0.0619	0.0146*	−0.0236	−2.3563
EM_1 = 0 Firm-Years								
DTE	1,644	0.0005	0.0128	0.0717	0.0055	0.0000	−0.0030	−0.0952
TAcc	1,644	−0.0500	0.1064	0.6549	−0.0050	−0.0446	−0.0857	−0.8689
AbAccMJ	1,644	−0.0110	0.1749	0.9975	0.0501	−0.0127	−0.0791	−0.8700
AbAccFL	1,295	−0.0071	0.1518	0.9801	0.0483	−0.0113	−0.0710	−0.8432
ΔCFO	1,644	−0.0008	0.1432	0.7247	0.0467	0.0043	−0.0406	−1.4483

Panel B: Earnings Management to Avoid a Loss Samples: Zero and Slightly Positive Earnings ($EM_2 = 1$) vs. Slightly Negative Earnings ($EM_2 = 0$), where $EM_2 = 1$ Firm-Years Have Scaled Earnings [(NI_t/MVE_{t-1})] of at Least 0 and Less Than 0.02 and $EM_2 = 0$ Firm-Years Have Scaled Earnings of at Least −0.02 and Less Than 0

					Percentiles			
EM_2 = 1 Firm-Years	**n**	**Mean**	**Std. Deviation**	**Maximum**	**75th**	**50th**	**25th**	**Minimum**
DTE	1,794	−0.0027*	0.0150	0.0734	0.0029	0.0000*	−0.0065	−0.0868
TAcc	1,794	−0.0332*	0.1162	0.8351	0.0155	−0.0430*	−0.0935	−0.8155
AbAccMJ	1,794	0.0049	0.1800	0.9735	0.0704	−0.0146	−0.0774	−0.9678
AbAccFL	1,473	0.0047	0.1742	0.9970	0.0679	−0.0080	−0.0771	−0.9958
CFO	1,794	0.0553*	0.1182	0.8446	0.1132	0.0614*	0.0024	−0.6767

(*continued on next page*)

TABLE 1 (continued)

$EM_2 = 0$ Firm-years	n	Mean	Std. Deviation	Maximum	Percentiles 75th	Percentiles 50th	Percentiles 25th	Minimum
DTE	991	−0.0056	0.0160	0.0685	0.0000	0.0000	−0.0087	−0.1226
TAcc	991	−0.0530	0.1169	0.7243	0.0008	−0.0520	−0.1041	−0.6425
AbAccMJ	991	−0.0053	0.1863	0.9372	0.0639	−0.0228	−0.0905	−0.9824
AbAccFL	781	−0.0008	0.1689	0.9717	0.0591	−0.0081	−0.0824	−0.7466
CFO	991	0.0183	0.1436	0.5179	0.0864	0.0331	−0.0277	−1.2737

Panel C: Earnings Management to Avoid Failing to Meet or Beat Analysts' Forecasts Samples: Zero or Slightly Positive Forecast Errors (EM_3 = 1) vs. Slightly Negative Forecast Errors (EM_3 = 0), where EM_3 = 1 Firm-Years Have Analyst Forecast Errors (Actual EPS less Mean Consensus Forecast) at Least 0 and Less Than or Equal to .01 and EM_3 = 0 Firm-Years Have Analyst Forecast Errors at Least −.01 and Less Than 0

$EM_3 = 1$ Firm-Years	n	Mean	Std. Deviation	Maximum	Percentiles 75th	Percentiles 50th	Percentiles 25th	Minimum
DTE	2,021	0.0007	0.0153	0.0868	0.0080	0.0003	−0.0046	−0.0917
TAcc	2,021	−0.0285	0.1018	0.8914	0.0115	−0.0358	−0.0794	−0.4573
AbAccMJ	2,021	−0.0060	0.1598	0.8732	0.0563	−0.0100	−0.0713	−0.9877
AbAccFL	1,743	−0.0013	0.1360	0.8366	0.0574	−0.0073	−0.0636	−0.8554
ΔCFO	2,021	0.0204	0.1431	0.8485	0.0731	0.0175	−0.0271	−2.2497
$EM_3 = 0$ Firm-Years								
DTE	509	0.0002	0.0151	0.0497	0.0070	0.0003	−0.0036	−0.0916
TAcc	509	−0.0358	0.1073	0.7730	0.0039	−0.0388	−0.0773	−0.5828
AbAccMJ	509	0.0001	0.1593	0.7799	0.0545	−0.0104	−0.0695	−0.5523
AbAccFL	436	0.0031	0.1326	0.8833	0.0487	−0.0089	−0.0639	−0.4766
ΔCFO	509	0.0027	0.2509	1.4256	0.0605	0.0120	−0.0323	−4.1109

(continued on next page)

TABLE 1 (continued)

* Mean (or median) $EM_t = 1$ variable is significantly different from mean (or median) $EM_t = 0$ variable in two-sided tests of means, at the <0.10 significance level.

Variable definitions:

DTE = deferred tax expense (annual Compustat data item #50), scaled by total assets (annual Compustat data item #6) at the end of year $t-1$;

TAcc = total accruals, scaled by total assets (annual Compustat data item #6) at the end of year $t-1$, is computed as *EBEI* − *CFO*, where *EBEI* is income before extraordinary items (annual Compustat data item #123), and *CFO* is cash flows from operations (computed as annual Compustat data items #308 − #124);

AbAccMJ = abnormal accruals computed using the modified Jones model (Dechow et al. 1995). It is calculated as the difference between *TAcc* and normal accruals. Modified Jones model normal accruals are estimated as: $TAcc_{it} = \alpha + \beta_1(\Delta Sales_{it} - \Delta REC_{it}) + \beta_2 PPE_{it}$, where $\Delta Sales$ is the change in the firm's sales (annual Compustat data item #12) from year $t-1$ to year t, ΔREC is the change in accounts receivable from operating activities (annual Compustat data item #302) from year $t-1$ to year t, and *PPE* is gross property, plant, and equipment (annual Compustat data item #7). All variables are scaled by total assets at the end of year $t-1$;

AbAccFL = abnormal accruals computed using the forward-looking Jones model (Dechow et al. 2002). It is calculated as the difference between *TAcc* and forward-looking normal accruals. Forward-looking model normal accruals are estimated as: $TAcc_{it} = \alpha + \beta_1(\Delta Sales_{it} - (1-k)\Delta REC_{it}) + \beta_2 PPE_{it} + \beta_3 TAcc_{it-1} + \beta_4 GR\_Sales_{t+1}$, where k is the slope coefficient from a regression of ΔREC_{it} on $\Delta Sales_{it}$, $TAcc_{it-1}$ is total accruals from year $t-1$, scaled by total assets as of year $t-2$, and GR_Sales_{t+1} is the change in sales from year t to $t+1$, scaled by year t sales;

CFO = *CFO* from year t, scaled by total assets as of the end of year $t-1$;

ΔCFO = the change in *CFO* from year $t-1$ to year t, scaled by total assets as of the end of year $t-1$;

EM_1 = an indicator variable, 1 if the change in net income (annual Compustat data item #172) from year $t-1$ to year t divided by the market value of equity at the end of year $t-2$ (annual Compustat data item #25 × #199) is ≥ 0 and <0.01, and $EM_1 = 0$ if the change in net income is ≥ -0.01 and <0;

EM_2 = an indicator variable, 1 if net income (annual Compustat data item #172) in year t divided by the market value of equity at the end of year $t-1$ (annual Compustat data item #25 × #199) is ≥ 0 and <0.02, and $EM_2 = 0$ if net income is ≥ -0.02 and <0; and

EM_3 = an indicator variable, 1 if the scaled forecast error (actual EPS less mean analysts' forecast) is ≥ 0 and ≤ 0.01, and $EM_3 = 0$ if the scaled forecast error is $\geq -.01$ and <0.

Firm-years are from the period 1994–2000, except for *AbAccFL*, where observations are from 1994–1999. The samples are trimmed at the 1st and 99th percentiles based upon *DTE*, *TAcc*, and *AbAccMJ*, or *AbAccFL*.

Descriptive statistics for our analysis of earnings levels appear in Panel B of Table 1. Consistent with *DTE* identifying earnings management activity to avoid a loss, the *DTE* mean of −0.0027 for the earnings interval of 0 to less than 0.02 of market value of equity is significantly greater (i.e., less negative) than the *DTE* mean of −0.0056 for the just-missed sample. (The median is also higher in the *EM* = 1 sample.) The negative mean *DTE*s indicate an average deferred tax *benefit*, which implies that the average firm in both earnings levels samples reports taxable income higher than book income. Average *TAcc* is also greater (i.e., less negative) in the *EM* = 1 sample, although neither the mean nor median for *AbAccMJ* or for *AbAccFL* differ between the samples.

Panel C of Table 1 reveals no significant differences in mean and median *DTE* between the *EM* = 1 and *EM* = 0 samples in the analysts' forecast setting. Means and medians for *TAcc* and the abnormal accrual variables also are not significantly different. These results are inconsistent with any of the measures detecting earnings management to meet or slightly beat analysts' forecasts.

Untabulated results indicate that across the three settings we consider, there are reliably positive correlations between change in net income and change in operating cash flows, consistent with Dechow (1994), and reliably negative correlations between total accruals and cash flows from operations, consistent with Sloan (1996). We find small but significantly positive correlations between *EM* and *DTE* in the earnings change and earnings levels settings. *EM* and *TAcc* are significantly positively associated in the earnings change and earnings levels settings, while correlations between *EM* and the abnormal accrual measures are insignificant. Additionally, *DTE* and the three accrual metrics generally are uncorrelated, which is consistent with Hanlon (2002) and suggests that multicollinearity between *DTE* and the accrual measures is not an issue. This low correlation between *DTE* and the accrual metrics also suggests that these measures may be proxying for different aspects of managerial discretion over accounting choices.

Primary Results

Our primary results provide evidence concerning the incremental usefulness of deferred tax expense *vis-à-vis* each accrual measure in detecting earnings management to avoid an earnings decline, to avoid a loss, and to avoid failing to meet or beat analysts' forecasts. Table 2 presents the results of comparing deferred tax expense (*DTE*) with total accruals (*TAcc*) across the three earnings management settings, Table 3 compares *DTE* with *AbAccMJ*, and Table 4 compares *DTE* and *AbAccFL*. Indicated p-values are one-tailed.

Deferred Tax Expense vs. Total Accruals

Table 2 reports the results of estimating probit models with both *DTE* and *TAcc* as test variables.[19] The first set of results is shown in the left-hand pair of columns labeled "Earnings Target 1: Scaled Earnings Changes" and concerns earnings management to avoid an earnings decline. The middle set of columns, labeled "Earnings Target 2: Scaled Earnings," presents the results for earnings management to avoid a loss, and the right-hand set of columns, labeled "Earnings Target 3: Analysts' Forecasts," displays the results related to earnings management to avoid failing to meet or beat analysts' earnings forecasts.

---

[19] We also estimate separate probit models with *DTE* as the only test variable in the model, and with *TAcc* (or *AbAccMJ* or *AbAccFl*) as the only test variable. In all settings, the coefficient significance levels for the individual test variables are qualitatively identical to those when both test variables (*DTE* and the accrual metric) are included in the model.

TABLE 2

Results of Probit Regressions for Three Earnings Targets: Comparison of Deferred Tax Expense (*DTE*) to Total Accruals (*TAcc*)

	Earnings Target 1: Scaled Earnings Changes		Earnings Target 2: Scaled Earnings		Earnings Target 3: Analysts' Forecast	
	n = 4,139[a]	Prob > χ^2 p-value	n = 2,785[a]	Prob > χ^2 p-value	n = 2,530[a]	Prob > χ^2 p-value
Intercept	0.3241	<0.0001	0.4807	<0.0001	0.9664	<0.0001
DTE	3.7785	0.0153	5.4334	0.0129	0.7243	0.7033
TAcc	0.9583	<0.0001	41.6193	<0.0001	0.6160	0.0422
ΔCFO	1.0961	<0.0001			0.4259	0.0206
CFO			41.1719	<0.0001		
Log Likelihood	−2,752		−1,157		−1,261	

Earnings Target 1: Scaled Earnings Changes. Earnings management to avoid an earnings decline, where $EM_1 = 1$ firm-years have scaled earnings changes $[(NI_t - NI_{t-1})/MVE_{t-1}]$ of at least 0 and less than 0.01 and $EM_1 = 0$ firm-years have scaled earnings changes of at least −0.01 and less than 0.

Earnings Target 2: Scaled Earnings. Earnings management to avoid a loss, where $EM_2 = 1$ firm-years have scaled earnings $[NI_t/MVE_t-1]$ of at least 0 and less than 0.02 and $EM_2 = 0$ firm-years have scaled earnings of at least −0.02 and less than 0.

Earnings Target 3: Analysts' Forecasts. Earnings management to avoid failing to meet or beat analysts' forecasts, where $EM_3 = 1$ firm-years have analyst forecast errors (actual EPS less mean consensus forecast) at least 0 and less than or equal to .01 and $EM_3 = 0$ firm-years have analyst forecast errors at least −.01 and less than 0.

[a] Sample trimmed based upon *DTE*, *TAcc*, and *AbAccMJ*.

Industry dummy variable results not shown.

See Table 1 for variable definitions.

With regard to the first set of results, the coefficient on *DTE* is positive (3.7785) and is significant ($p = 0.0153$), suggesting that deferred tax expense is incrementally useful in detecting earnings management to avoid an earnings decline after controlling for total accruals, changes in operating cash flows, and industry. The coefficient on *TAcc* equals 0.9583, and is also significant ($p < 0.0001$), indicating that total accruals is also incrementally useful in detecting earnings management in this setting. As expected, ΔCFO is positive and significant.

The results reported in the middle column Table 2 are similar to the first set. The coefficient on *DTE* is positive (5.4334) and significant ($p = 0.0129$), as is the coefficient on *TAcc* (41.6193, $p < 0.0001$). Hence, deferred tax expense is incrementally useful beyond total accruals in identifying earnings management to avoid a loss. Likewise, *TAcc* is incrementally useful beyond DTE in this setting. The coefficient on *CFO* is also reliably positive.

In the third column of results in Table 2, *TAcc* has a significant coefficient ($p = 0.0422$), whereas *DTE* does not, and thus only total accruals is incrementally useful in detecting earnings management to meet analysts' forecasts. Recall that prior research on earnings management in this setting provides conflicting evidence, and there is evidence that managerial guidance of analysts' forecasts is perhaps more important in this setting.

Deferred Tax Expense vs. Abnormal Accruals

We report the results of the comparison of deferred tax expense and abnormal accruals derived from the modified Jones model in Table 3. The coefficient on *DTE* is positive and significant when considering both scaled earnings changes (3.6868, $p = 0.0178$) and scaled

TABLE 3
Results of Probit Regressions for Three Earnings Targets: Comparison of Deferred Tax Expense (*DTE*) to Modified-Jones Abnormal Accruals (*AbAccMJ*)

	Earnings Target 1: Scaled Earnings Changes		**Earnings Target 2: Scaled Earnings**		**Earnings Target 3: Analysts' Forecast**	
	n = 4,139[a]	**Prob > χ^2 p-value**	**n = 2,785**[a]	**Prob > χ^2 p-value**	**n = 2,530**[a]	**Prob > χ^2 p-value**
Intercept	0.2571	<0.0001	0.2810	<0.0001	0.9401	<0.0001
DTE	3.6868	0.0178	8.8055	<0.0001	0.8058	0.6720
AbAccMJ	0.1423	0.2506	0.8854	<0.0001	−0.0730	0.6925
ΔCFO	0.8344	<0.0001			0.2932	0.0711
CFO			2.1777	<0.0001		
Log Likelihood	−2,761		−1,752		−1,263	

Earnings Target 1: Scaled Earnings Changes. Earnings management to avoid an earnings decline, where $EM_1 = 1$ firm-years have scaled earnings changes $[(NI_t - NI_{t-1})/MVE_{t-1}]$ of at least 0 and less than 0.01 and $EM_1 = 0$ firm-years have scaled earnings changes of at least −0.01 and less than 0.

Earnings Target 2: Scaled Earnings. Earnings management to avoid a loss, where $EM_2 = 1$ firm-years have scaled earnings $[NI_t/MVE_{t-1}]$ of at least 0 and less than 0.02 and $EM_2 = 0$ firm-years have scaled earnings of at least −0.02 and less than 0.

Earnings Target 3: Analysts' Forecasts. Earnings management to avoid failing to meet or beat analysts' forecasts, where $EM_3 = 1$ firm-years have analyst forecast errors (actual EPS less mean consensus forecast) at least 0 and less than or equal to .01 and $EM_3 = 0$ firm-years have analyst forecast errors at least −.01 and less than 0.

[a] Sample trimmed based upon *DTE*, *TAcc*, and *AbAccMJ*.

Industry dummy variable results not shown.

See Table 1 for variable definitions.

earnings levels (8.8055, $p < 0.0001$), but is again insignificant for the analysis of analysts' forecast errors. The coefficient on *AbAccMJ* is only significant in the setting in which firms manage earnings to avoid a loss ("Earnings Target 2" in Table 3). Thus, *DTE* is incrementally useful to modified Jones model abnormal accruals in detecting earnings management to avoid reporting an earnings decline and to avoid a loss, whereas modified Jones model abnormal accruals are incrementally useful only in the latter setting.

Table 4 presents the comparison of *DTE* and abnormal accruals estimated using the forward-looking model. The coefficient on *DTE* only approaches statistical significance ($p = 0.1157$) in detecting earnings management to avoid an earnings decline, whereas the coefficient on *AbAccFL* is positive and significant ($p = 0.0002$). The coefficients on *DTE* and *AbAccFL* are both significantly positive in the analysis of earnings levels ($p < 0.0001$ and $p = 0.0008$, respectively), while both variables are insignificant in the analysts' forecast context. Thus, *DTE* is incrementally useful to forward-looking abnormal accruals in the earnings levels setting, while forward-looking abnormal accruals, which reflects current and future period information, is incrementally useful in both the earnings changes and levels settings.

Summary of Primary Results

Overall, the results are consistent with H1 and H2, but not H3. Deferred tax expense is generally incrementally useful to the various accrual-based measures in detecting earnings management to avoid an earnings decline (H1) and to avoid reporting a loss (H2), but not in detecting earnings management with regard to avoiding failing to meet or beat analysts'

TABLE 4

Results of Probit Regressions for Three Earnings Targets: Comparison of Deferred Tax Expense (*DTE*) to Forward-Looking Abnormal Accruals (*AbAccFL*)

	Earnings Target 1: Scaled Earnings Changes		**Earnings Target 2: Scaled Earnings**		**Earnings Target 3: Analysts' Forecast**	
	n = 3,352^a	**Prob > χ^2 p-value**	**n = 2,254**^{a}	**Prob > χ^2 p-value**	**n = 2,279**^{a}	**Prob > χ^2 p-value**
Intercept	0.2771	<0.0001	0.3496	<0.0001	1.0197	<0.0001
DTE	2.7257	0.1157	10.2743	<0.0001	0.3648	0.8630
AbAccFL	0.6395	0.0002	0.6174	0.0008	−0.0039	0.9868
ΔCFO	1.2130	<0.0001			0.2785	0.1097
CFO			1.6944	<0.0001		
Log Likelihood	−2,212		−1,417		−1,083	

Earnings Target 1: Scaled Earnings Changes. Earnings management to avoid an earnings decline, where $EM_1 = 1$ firm-years have scaled earnings changes $[(NI_t - NI_{t-1})/MVE_{t-1}]$ of at least 0 and less than 0.01 and $EM_1 = 0$ firm-years have scaled earnings changes of at least −0.01 and less than 0.

Earnings Target 2: Scaled Earnings. Earnings management to avoid a loss, where $EM_2 = 1$ firm-years have scaled earnings $[NI_t/MVE_{t-1}]$ of at least 0 and less than 0.02 and $EM_2 = 0$ firm-years have scaled earnings of at least −0.02 and less than 0.

Earnings Target 3: Analysts' Forecasts. Earnings management to avoid failing to meet or beat analysts' forecasts, where $EM_3 = 1$ firm-years have analyst forecast errors (actual EPS less mean consensus forecast) at least 0 and less than or equal to .01 and $EM_3 = 0$ firm-years have analyst forecast errors at least −.01 and less than 0.

[a] Sample trimmed based upon *DTE*, *TAcc*, and *AbAccFL*.

Industry dummy variable results not shown.

See Table 1 for variable definitions.

forecasts (H3). Abnormal accruals derived from either the modified Jones model or the forward-looking model are sometimes incrementally useful to DTE in detecting earnings management. Somewhat surprisingly, total accruals is incrementally useful to DTE in detecting earnings management in all three settings we consider.

Supplemental Analysis

We perform an additional analysis in which we consider the relative, rather than incremental, usefulness of deferred tax expense versus the accrual-based measures in detecting earnings management by assessing their ability to classify firm-years as *EM* = 1 or *EM* = 0. More specifically, we perform a Receiver Operator Characteristic (ROC) analysis (see Sprinkle and Tubbs 1998), which Hosmer and Lemeshow (2000) argue is appropriate for this purpose when using logit-type models.[20] Compared to a naïve classification percentage, ROC analysis provides a more complete description of classification accuracy by plotting the probability of detecting a true positive and a false positive using the entire distribution of the test diagnostic (e.g., *DTE* and the accrual-based metrics) as potential cut points; i.e., observations above the test diagnostic's cut point are classified as positives (e.g., *EM* = 1) and those below as negatives (e.g., *EM* = 0). For the settings we consider, the area under the ROC curve is determined by randomly selecting pairs of firm-years, an *EM* = 1 and *EM* = 0, and testing to determine if the respective earnings management measure (e.g.,

---

[20] We do not perform the Vuong test to compare the relative ability of *DTE* and the accrual-based measures to detect earnings management because this test relies on OLS estimation, which is inappropriate because our dependent variable is a dichotomous variable.

DTE) is higher for the $EM = 1$ firm-year relative to the $EM = 0$ firm-year. The area under the ROC curve equals the percentage of randomly drawn pairs for which the earnings management measure is in fact higher for the $EM = 1$ observations.

In Table 5, we summarize the results of the ROC analysis. The area under the ROC curve is not significantly different between *DTE* and each of the accrual-based measures in the scaled earnings changes and analysts' forecast settings, but it is significantly greater for *DTE* than for each of the accrual-based measures in the scaled earnings levels setting. Thus, *DTE* is relatively more useful than total accruals and both modified Jones model and forward-looking model abnormal accruals in accurately classifying firm-years as earnings management or non-earnings management firm-years with respect to avoiding a loss. In the setting in which firms manage earnings to avoid an earnings decline or to avoid failing to meet or beat the consensus analysts' forecast, neither *DTE* nor any accrual-based measure is relatively more useful in classifying firm-years. We acknowledge that the areas under the ROC curve are considered poor (Hosmer and Lemeshow 2000). This result is analogous to studies that document a relation but do so with models having low R^2s.

Sensitivity Analysis

We assess the sensitivity of the primary results in a number of additional ways (details not shown). First, we use logit and OLS regressions in place of probit and obtain similar results. Second, we augment the probit models for earnings changes and for analysts' forecasts with the level of operating cash flows (and with ΔCFO for the earnings levels case) to further control for current performance. Our inferences remain unchanged. This is also the case when we include the square of the cash flow variable to control for possible

TABLE 5
Results of Receiver Operator Characteristic (ROC) Tests

	Earnings Target 1: Scaled Earnings Changes		Earnings Target 2: Scaled Earnings		Earnings Target 3: Analysts' Forecasts	
	Area under ROC Curve	p-value for difference[a]	Area under ROC Curve	p-value for difference[a]	Area under ROC Curve	p-value for difference[a]
n	4,139		2,785		2,530	
DTE	0.5266		0.5709		0.4985	
TAcc	0.5195	0.5916	0.5446	0.0856	0.5107	0.5535
AbAccMJ	0.5064	0.1162	0.5188	0.0009	0.4964	0.9158
n	3,352		2,254		2,179	
DTE	0.5223		0.5819		0.5014	
AbAccFL	0.5135	0.5392	0.5146	0.0001	0.4998	0.9400

[a] Two-sided p-value for a χ^2 test of the difference between the areas under the ROC curves for *DTE* vs. the indicated accrual measure.

Earnings Target 1: Scaled Earnings Changes. Earnings management to avoid an earnings decline, where $EM_1 = 1$ firm-years have scaled earnings changes $[(NI_t - NI_{t-1})/MVE_{t-1}]$ of at least 0 and less than 0.01 and $EM_1 = 0$ firm-years have scaled earnings changes of at least −0.01 and less than 0.

Earnings Target 2: Scaled Earnings. Earnings management to avoid a loss, where $EM_2 = 1$ firm-years have scaled earnings $[NI_t/MVE_{t-1}]$ of at least 0 and less than 0.02 and $EM_2 = 0$ firm-years have scaled earnings of at least −0.02 and less than 0.

Earnings Target 3: Analysts' Forecasts. Earnings management to avoid failing to meet or beat analysts' forecasts, where $EM_3 = 1$ firm-years have analyst forecast errors (actual EPS less mean consensus forecast) at least 0 and less than or equal to .01 and $EM_3 = 0$ firm-years have analyst forecast errors at least −.01 and less than 0.

See Table 1 for variable definitions.

nonlinearity in the relation between *EM* and (change in) operating cash flows. Third, rather than pooling across firm-years, we estimate separate probit regressions for each of the seven sample years. In the earnings change analysis the mean estimated coefficient on *DTE* is 4.97 (Fama-MacBeth t-statistic = 2.07, p-value < 0.03), and in the earnings level analysis the mean *DTE* coefficient across the seven annual regressions is 6.27 (Fama-MacBeth t-statistic = 4.47, p-value < 0.0001).

Fourth, we augment the probit regressions with three additional control variables. The first two control variables are asset growth and revenue growth, each measured over the prior three years. We include these variables to control for the possibility that *DTE* is proxying for growth (e.g., Manzon and Plesko 2002) and that *EM* = 1 firms are simply high-growth firms relative to *EM* = 0 firms. The third control variable is average return on assets over the prior three years and serves as an additional control for performance (i.e., beyond [changes in] operating cash flows), which may affect the ability of accrual variables to detect earnings management (McNichols 2000). Including these control variables in the probit models does not eliminate the significance of *DTE.*

Fifth, we further investigate whether firm performance drives our result that *DTE* is incrementally useful to the accrual-based metrics in detecting earnings management to avoid an earnings decline. It is possible that high *DTE* is a proxy for high performance and that *EM* = 1 firms are simply higher performing firms than *EM* = 0 firms. We define performance alternatively as pretax and after-tax return on lagged total assets (*ROA*), further define high-performance firm-years as those in the upper third of the distribution of the performance variable, and drop firm-years in the middle third of the distribution to focus on performance extremes (e.g., McNichols 2000). We then regress *EM* on *DTE*, the accrual variable, the dummy variable indicating high performance, and multiplicative terms interacting the high-performance dummy variable with *DTE* and the particular accrual variable, respectively. Our primary results are unaffected and the coefficients on the *DTE* × *High ROA* variables are insignificant. However, controlling for firm performance makes the estimated coefficients on the accrual measures insignificant. Thus, while firm performance does not drive the significance of the *DTE* results, controlling for firm performance takes away the significance, and thus the incremental usefulness, of the accrual metrics' abilities to detect earnings management in this setting (Kothari et al. 2002).[21]

Sixth, because deferred taxes and accruals reverse in the subsequent year(s), we investigate whether such reversals affect our results. We include EM_{t-1} and define it to equal 1 if a firm's earnings change (or earnings level) fell into the meet or just beat interval in the previous period, and 0 otherwise. We also interact EM_{t-1} with year *t*'s *DTE* and accrual metric. If reversals affect our results, then the coefficient on $DTE_t \times EM_{t-1}$ will be negative in both the avoid an earnings decline and avoid a loss settings. Our primary results remain unchanged and the coefficient on the interaction of DTE_t and EM_{t-1} is insignificant. We interpret these results as evidence that the prior year's earnings managers are not experiencing reversals of temporary book-tax differences that are driving our primary findings.

Seventh, we perform an analysis to investigate whether firms that manage earnings in year *t* have higher *DTE* only in year *t*, and not in surrounding years. Specifically, we regress this period's *EM* on last period's *DTE* and *TAcc* (as well as operating cash flows and industry), and separately on next period's *DTE* and *TAcc*. In both the earnings decline and loss settings, the results indicate that the coefficients on the lead and the lag values of *DTE* are not significantly positive. This evidence is consistent with *DTE* being positively related to *EM* only in the year a firm is suspected of being an earnings manager.

---

[21] We cannot perform this analysis in the earnings levels setting since the performance variables and *EM* to avoid a loss are almost perfectly correlated. That is, because the *EM* = 1 vs. *EM* = 0 classification depends solely on net income, a performance measure that includes income produces a tautological relation with *EM*.

Finally, we investigate whether tax-planning actions that lower taxable income relative to book income can explain the positive relation between *DTE* and *EM*, as opposed to attributing this positive relation to firms managing book income upward. We do this in two ways. First, we compute mean and median current effective income tax rates (*CurrETR*), defined as current tax expense divided by pretax book income (annual Compustat data items [(#63 + #64)/#170]). We find that on average these rates are higher for *EM* = 1 firm-years in both the avoid an earnings decline and avoid a loss settings. Higher current effective tax rates for EM = 1 firm-years suggests it is unlikely that the higher average *DTE* for these firm-years results from tax-planning strategies that reduce taxable income relative to book income, since such strategies would reduce current tax expense and current effective tax rates.

Second, we include in our probit regressions a dummy variable for low current tax expense, which equals 1 if a firm-year falls in the bottom third of the distribution of current tax expense scaled by total assets. We also interact that dummy variable with *DTE* and with the accrual variable. We re-estimate the augmented models and find that our primary results are unaffected, and the coefficient on *DTE* × *Low Current Tax Expense* is not significant. We also perform a similar analysis by deleting firm-years in the bottom third of the distribution of current tax expense scaled by total assets and our primary results remain unchanged. We conclude that low current tax expense firm-years (i.e., those that engage in tax planning to lower taxable income relative to book income) do not drive the positive relation between *DTE* and *EM*.

Overall, the sensitivity analyses reinforce the results that are consistent with the basic hypotheses that deferred tax expense is incrementally useful to the accrual metrics in detecting earnings management to avoid an earnings decline and to avoid a loss.

V. CONCLUSIONS

We investigate the incremental usefulness of deferred tax expense in detecting earnings management. Because earnings management is accomplished using managerial discretion, and because managers generally have more discretion under GAAP than under tax rules, we expect that managers will manage earnings upward by exploiting their discretion under GAAP, and will presumably do so in a manner that does not increase current income taxes payable. If so, such earnings management behavior will generate temporary book-tax differences that lead to higher deferred tax expense. Building on evidence of earnings management in Burgstahler and Dichev (1997), Degeorge et al. (1999), and Mills and Newberry (2001), and on evidence of significant measurement error in accrual measures in Guay et al. (1996) and Bernard and Skinner (1996), we evaluate the incremental abilities of deferred tax expense and three accrual-based measures to detect earnings management to meet or slightly beat three earnings targets: last year's earnings, zero earnings, and consensus analysts' forecasts.

The results are consistent with the incremental usefulness of deferred tax expense in detecting earnings management. Deferred tax expense is incrementally useful to total accruals and modified Jones model abnormal accruals in detecting earnings management to avoid an earnings decline, and is incrementally useful beyond total accruals, modified Jones model abnormal accruals, and forward-looking abnormal accruals in detecting earnings management to avoid a loss. We also find that *DTE* is significantly more accurate than the three accrual measures in classifying firm-years as earnings management and non-earnings management firm-years with regard to avoiding a loss. *DTE* is not incrementally useful in detecting earnings management to avoid failing to meet or beat analysts' forecasts.

Our results add to recent findings that indicate a relation between book and tax reporting and firms' incentives to engage in earnings management activities (Mills and Newberry 2001). The evidence in our paper suggests that *DTE* can supplement accrual measures in detecting earnings management to avoid an earnings decline and to avoid a loss. Surprisingly, our results suggest that total accruals is incrementally useful in detecting earnings management activities in the three settings we consider, while the performance of abnormal accrual measures is mixed. Further, there is evidence that firm performance can adversely affect the usefulness of the accrual measures in detecting earnings management, whereas that does not appear to be the case for deferred tax expense.

One limitation of our study is that our analysis is restricted to the period in which SFAS No. 109 has been in effect. Another limitation is that we have not incorporated managerial guidance into our investigation of the ability to detect earnings management to avoid failing to meet or beat analysts' earnings forecasts; we leave that for future research. Future research might also fruitfully consider (1) modeling the determinants of deferred tax expense to discover whether *DTE* itself can be usefully decomposed, (2) identifying the components of *DTE* that reflect most earnings management activity, and (3) examining the usefulness of DTE and accrual variables to detect earnings management in quarterly data and in other settings.

REFERENCES

Baber, W., and S-H. Kang. 2002. Is meeting the consensus EPS good news or bad news? Stock splits and the accuracy of analysts forecast data. Working paper, The George Washington University.

Barth, M., J. Elliott, and M. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37 (2): 387–413.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (2): 173–204.

Bauman, C., M. Bauman, and R. Halsey. 2001. Do firms use the deferred tax asset valuation allowance to manage earnings? *The Journal of the American Taxation Association* 23 (Supplement): 27–48.

Beatty, A., B. Ke, and K. Petroni. 2000. Earnings management to avoid earnings declines across publicly and privately held banks. *The Accounting Review* 77 (3): 547–570.

Bernard, V., and D. Skinner. 1996. What motivates managers' choice of discretionary accruals? *Journal of Accounting and Economics* 22 (1-3): 313–325.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (1): 99–126.

———, and M. Eames. 2002. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. Working paper, University of Washington.

———, W. B. Elliott, and M. Hanlon. 2002. How firms avoid losses: Evidence of use of the net deferred tax asset account. Working paper, University of Washington.

DeAngelo, L. 1986. Management buyouts of public stockholders. *The Accounting Review* 61 (3): 400–420.

Dechow, P. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18 (1): 3–42.

———, R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, S. Richardson, and A. I. Tuna. 2003. Why are earnings kinky? An examination of the earnings management explanation. *Review of Accounting Studies* (forthcoming).

DeFond, M., and J. Jiambalvo. 1994. Debt covenant violations and manipulation of accruals. *Journal of Accounting and Economics* 17 (1-2): 145–176.

———, and K. R. Subramanyam. 1998. Auditor changes and discretionary accruals. *Journal of Accounting and Economics* 25(1): 35–67.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72 (1): 1–33.

Dhaliwal, D., C. Gleason, and L. Mills. 2002. Last chance earnings management: Using the tax expense to achieve earnings targets. Working paper, The University of Arizona.

Erickson, M., M. Hanlon, and E. Maydew. 2002. How much will firms pay for earnings that do not exist? Evidence of taxes paid on allegedly fraudulent earnings. Working paper, University of Chicago.

Financial Accounting Standards Board (FASB). 1992. *Accounting for Income Taxes.* Statement of Financial Accounting Standards No. 109. Stamford, CT: FASB.

Guay, W., S. P. Kothari, and R. Watts. 1996. A market-based evaluation of discretionary accrual models. *Journal of Accounting Research* 34 (Supplement): 83–105.

Hanlon, M. 2002. The persistence and pricing of earnings, accruals, and cash flows when firms have large book-tax differences. Ph.D. dissertation, University of Washington.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (1-3): 85–107.

Hosmer, D., and S. Lemeshow. 2000. *Applied Logistic Regressions.* 2nd edition. New York, NY: John Wiley.

Hribar, P., and D. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (1): 105–134.

Hunt, A., S. Moyer, and T. Shevlin. 1996. Managing interacting accounting measures to meet multiple objectives: A study of LIFO firms. *Journal of Accounting and Economics* 21 (3): 339–374.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.

Joos, P., J. Pratt, and S. D. Young. 2002. Using deferred taxes to infer the quality of accruals. Working paper, Massachusetts Institute of Technology.

Kasznik, R., and M. McNichols. 2002. Does meeting expectations matter? Evidence from analysts forecast revisions and share prices. *Journal of Accounting Research* 40 (3): 727–759.

Kothari, S. P., A. Leone, and C. Wasley. 2002. Performance matched discretionary accrual measures. Working paper, Massachusetts Institute of Technology.

Manzon, G., and G. Plesko. 2002. The relation between financial and tax reporting measures of income. *Tax Law Review* 55 (2): 175–214.

Matsumoto, D. 2002. Management's incentives to avoid earnings surprises. *The Accounting Review* 77 (3): 483–514.

McNichols, M. 2000. Research design issues in earnings management studies. *Journal of Accounting and Public Policy* 19 (4–5): 313–345.

Miller, G., and D. Skinner. 1998. Determinants of the valuation allowance for deferred tax assets under SFAS No. 109. *The Accounting Review* 73 (2): 213–233.

Mills, L., and K. Newberry. 2001. The influence of tax and nontax costs on book-tax reporting differences: Public and private firms. *The Journal of the American Taxation Association* 23 (1): 1–19.

Myers, L., and D. Skinner. 2001. Earnings momentum and earnings management. Working paper, University of Michigan.

Payne, J., and W. Thomas. 2002. The implications of using stock-split adjusted I/B/E/S data in empirical research. Working paper, The University of Oklahoma.

Plesko, G. 2002. Reconciling corporations' book and taxable income, 1996–1998. *SOI Bulletin* (Spring): 111–132.

Schwartz, W. 2001. Managing earnings surprises. Ph.D. dissertation, The University of Iowa.

Schrand, C., and F. Wong. 2002. Earnings management using the valuation allowance for deferred tax assets under SFAS 109. Working paper, University of Pennsylvania.

Skinner, D., and R. Sloan. 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7 (2/3): 289–312.

Slemrod, J. 2002. Tax minimization and corporate responsibility. *96 Tax Notes* (Sept. 9): 1523–1529.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

Sprinkle, G., and R. Tubbs. 1998. The effects of audit risk and information importance on auditor memory during working paper review. *The Accounting Review* 73 (4): 475–502.

Yin, G. 2001. Getting serious about corporate tax shelters: Taking a lesson from history. *SMU Law Review* 54 (1): 209.

THE ACCOUNTING REVIEW
Vol. 78, No. 2
2003
pp. 523–553

Why Has the Contemporaneous Linear Returns-Earnings Relation Declined?

Stephen G. Ryan
Paul A. Zarowin
New York University

ABSTRACT: We investigate two explanations for the declining contemporaneous linear relation between annual stock returns and accounting earnings over the past 30 years: (1) earnings increasingly reflect news with a lag relative to stock prices and (2) earnings increasingly reflect good and bad news in an asymmetric fashion. We hypothesize and find that annual earnings have a weaker association with current price changes and a stronger association with lagged price changes over time. We hypothesize and find that annual earnings reflect current positive price changes less strongly and current negative price changes more strongly over time. We also find that asymmetry with respect to lagged price changes is increasingly important over time. Strikingly, we find that, since the mid-1980s, the aggregation of earnings over a four-year window increasingly does less to reduce the importance of lags and asymmetry.

Keywords: *capital markets; returns-earnings relation; earnings lags; earnings asymmetry.*

Data Availability: *The data used in this study are publicly available.*

I. INTRODUCTION

Prior research shows that the contemporaneous linear relation between annual stock returns and accounting earnings has declined over the past 30 years. For example, Francis and Schipper (1999), Lev and Zarowin (1999), and Ely and Waymire (1999) estimate regressions of annual (or 15-month) returns on the level and/or change in annual earnings deflated by beginning market value, and they find that the slope coefficients and R^{2}s in these regressions decrease strongly over time.[1] In this paper, we investigate two explanations for this declining relation: (1) earnings increasingly reflect news with a lag relative to stock prices (hereafter lags) and (2) earnings increasingly reflect good and bad news in an asymmetric fashion (hereafter asymmetry). These explanations interact, because if there is asymmetry with respect to the current price change, then there must be asymmetry

---

[1] Relatedly, Collins et al. (1997) estimate an equation with market value regressed on book value and earnings, and they find similar declining value relevance for earnings but not book value.

The authors thank three anonymous referees, Bill Beaver, Mary Ellen Carter, Craig Nichols, Gerry Salamon, Katherine Schipper, Jim Wahlen, and seminar participants at the 2001 Conference on Financial Economics and Accounting, Arizona State University, Indiana University, New York University, Rutgers University–Newark, and Temple University for useful comments.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted September 2002
Accepted November 2002

with respect to lagged price changes. This investigation is important because accounting policymakers and users of financial reports must identify the reasons for the declining returns-earnings relation to decide what, if anything, should be done to address the situation.

While prior researchers have examined lags and asymmetry, they either have studied each individually or have not focused on intertemporal changes, and so have not provided as full a story for the declining returns-earnings association as we do. Regarding lags, Kothari and Sloan (1992), Beaver and Ryan (1993), and Collins et al. (1994), among others, show that annual returns predict future earnings for at least three subsequent years. None of these papers examine intertemporal changes in lags or incorporate asymmetry. Regarding asymmetry, Basu (1997) and Givoly and Hayn (2000) show that, since the early 1970s, the association between contemporaneous returns and earnings is stronger when returns are negative than when they are positive, reflecting increasing accounting conservatism attributable at least in part to increasing legal liability. Neither of these papers examines lags or how they have changed over time.

Similar to Pope and Walker (1999), our empirical analysis is based on an equation that regresses annual earnings on current and three lagged annual price changes, with dummy intercept and slope coefficients for negative price changes, and with all variables deflated by price at the beginning of the current year. This equation captures earnings' association with price changes at up to a three-year lag, and it allows the responsiveness of earnings to current and lagged price changes to depend on the signs of the price changes. We estimate this equation and restricted versions of it (without the lagged price changes and/or dummy coefficients) for annual cross-sections from 1966 to 2000, in order to assess changes in the joint and incremental effects of lags and asymmetry over that period.

To motivate this regression equation and formulate precise hypothesis tests, we extend Ryan's (1995) model of accrual measurement (which includes lags) to include asymmetry. Our theoretical model is similar to those in Pope and Walker (1999) and Ball et al. (2000) but has an important difference: our model implies that the coefficient on a given price change is the portion of that price change that is reflected in current earnings. Moreover, the sum of the coefficients on the set of positive (or negative) price changes from 1 to s lags equals the portion of a price change of that sign that is cumulatively recognized in earnings from 1 to s years after the price change. Thus, the model allows us to interpret changes in (sums of) coefficients across time as measures of changes in earnings lags and asymmetry, and so eliminates the need to rely on the R^2 comparisons used in prior research. As Gu (2002) shows, intertemporal R^2 comparisons can be problematic due to differences in scale and other factors across samples, although Brown et al. (1999) find that this is mainly the case for levels models, not for price-deflated changes models such as ours.

We do analyze intertemporal changes in the *incremental* R^2s of less restricted regression equations to the most restricted equation, however. While these incremental R^2s are also sample-dependent, this dependence is much less of an issue than for the raw R^2s, since the R^2 for the most restricted equation constitutes a benchmark that incorporates sample changes unrelated to lags or asymmetry.

We find strongly increasing lags and asymmetry. Regarding lags, we find that earnings have a weaker association with current price changes and a stronger association with lagged price changes over time. Relatedly, we find that the incremental R^2s from including the lagged price changes in the regression equations rise strongly over time. Regarding asymmetry, we find that earnings reflect current positive price changes less strongly and current negative price changes more strongly over time. We also find that the importance of asymmetry with respect to lagged price changes increases over time. Relatedly, we find that the incremental R^2s from including the dummy coefficients in the regression equations rise strongly over time.

Our finding that annual earnings exhibit increasing lags and asymmetry raises the question of whether earnings measured over longer windows exhibit similar changes. Easton et al. (1992) show that the linear returns-earnings relation strengthens as returns and earnings are measured over longer windows, suggesting less lags and asymmetry than for annual earnings, while Basu (1997) shows that asymmetry lessens over longer windows. However, this prior research does not provide evidence on changes in lags and asymmetry for multiyear earnings over time. To provide such evidence, we develop and estimate regression equations with earnings aggregated over four years as the dependent variable. We find that, since the mid-1980s, an equation with four-year earnings that does not allow for lags and asymmetry has a sharply declining R^2, and the incremental R^2s from adding these factors rise dramatically. Thus, aggregating earnings over four years solves less of the decline in the contemporaneous linear returns-earnings association over time, because such aggregated earnings also exhibit increasing lags and asymmetry.

We conduct three analyses to probe our main results. First, increasing lags and asymmetry could be attributable to firms continuously in the sample or to the net addition to the sample of firms for which these factors are relatively important, and so we also conduct our analyses on a constant sample of firms. The constant sample also exhibits increasing lags and asymmetry, suggesting that the increasing lags and asymmetry that characterize the full sample are not due only to changes in sample composition over time.

Second, Pope and Walker (1999) find that the results of estimating regression equations similar to ours are sensitive to deflation by four-year rather than one-year lagged price. We explain this sensitivity by showing theoretically that deflation by s-year lagged price makes the coefficients on the 1 to $s-1$ lagged price changes depend on discount rates. Since discount rates trend down since the early 1980s, the effects of decreasing discount rates and increasing lags on these coefficients offset, making it difficult to observe increasing lags.

Third, and most important, lags and asymmetry could result from the nature of accounting rules, managerial reporting choices, non-earnings information, market inefficiency, or competition. While we largely leave the task of identifying the specific nature of these factors to future research, we provide preliminary evidence on the relative importance of accounting factors that are reflected in accruals vs. economic factors that are reflected in cash flow from operations (CFO). Specifically, we estimate our regression equations using CFO rather than earnings as the dependent variable, as in Basu (1997). We find that CFO exhibits lags and asymmetry at much lower levels that change less over time than does earnings, suggesting that accounting is the primary, but not sole, reason for increasing lags and asymmetry.

The remainder of this paper is as follows. Section II develops the theoretical model and regression equations. Section III develops the hypotheses. Section IV describes the sample and descriptive statistics. Section V reports the regression results for the full and constant samples using annual earnings as the dependent variable. Section VI develops and reports results for regressions using four-year earnings as the dependent variable. Section VII reports results using annual cash flow from operations as the dependent variable. Section VIII concludes.

II. THEORETICAL MODEL AND REGRESSION EQUATIONS

In this section, we extend the model of accrual measurement in Ryan (1995) in order to motivate the specification, interpretation, and testing of our regression equation. Ryan's

(1995) Equation (1) incorporates lags, modeling the book value of common equity as equal to the market value of common equity less a moving average process in current and prior shocks to market value.[2] We incorporate asymmetry into this model by allowing the coefficients to depend on the signs of the shocks to market value. Using this model, we derive an expression for earnings that leads to a regression equation with interpretable coefficients.

We model earnings measurement using three assumptions. First, we assume no arbitrage (NA):

$$MV_t - MV_{t-1} + D_t = rMV_{t-1} + \phi_t, \quad \text{with } E_{t-1}\{\phi_t\} = 0, \tag{NA}$$

where MV_t denotes market value of common shareholders equity at the end of year t, D_t denotes net dividends in year t, r denotes the cross-sectional constant cost of equity capital, and ϕ_t is a mean zero shock to market value in year t. Firm subscripts are suppressed on all variables.[3]

Second, generalizing Ryan's (1995) model as described above, we assume that the nature of accrual measurement (AM) implies that the book value of common equity at the end of year t, denoted BV_t, equals MV_t less a moving average process in the current and prior shocks, ϕ_{t-s}, where the coefficients in the process are functions F_s of ϕ_{t-s}; i.e., $F_s(\phi_{t-s})$ denotes the portion of ϕ_{t-s} that is not reflected in BV_t. As in Ryan (1995), to capture lags, we assume that for each realization of ϕ_t the value of $F_s(\phi_t)$ declines toward 0 with lag s, so that the shocks are gradually realized in book value.[4] Extending Ryan (1995), to capture asymmetry, we assume that the F functions take constant values for non-negative shocks and different constant values for negative shocks, where the dummy variable $D\phi_t$ takes a value of 1 if ϕ_1 is negative, and 0 otherwise:

$$\begin{aligned} BV_t &= MV_t - F_0(\phi_t)\phi_t - F_1(\phi_{t-1}) - F_2(\phi_{t-2})\phi_{t-2} - \dots, \\ &\quad \text{with } 1 \geq F_0(\phi_t) \geq F_1(\phi_t) \geq F_2(\phi_t) \geq \cdots \geq 0, \\ &\quad \text{and } F_s(\phi_t) = \begin{Bmatrix} F_{sp} \text{ if } D\phi_t = 0 \\ F_{sp} + F_{sn} \text{ if } D\phi_t = 1 \end{Bmatrix}. \end{aligned} \tag{AM}$$

Ryan (1995) analyzes a firm with assets with finite life K, which implies that $F_{K-1}(\phi_t) = 0$. Since intangibles such as goodwill may have an indefinite life, we allow the moving average process to have an indefinite length. Note that AM is silent as to whether accounting numbers convey ϕ_t to the market in year t or are simply associated with ϕ_t.

Third, we assume clean surplus accounting (CS), so that earnings, denoted X, equal the change in book value plus net dividends:

$$X_t = BV_t - BV_{t-1} + D_t. \tag{CS}$$

CS ensures that shocks are fully recognized in earnings over time, which enhances the interpretability of the (sums of) coefficients in our regression equations.

---

[2] Ryan's (1995) model has parallels to those in Pope and Walker (1999) and Ball et al. (2000), among others, but is preferable for our purposes because it provides greater interpretability of (sums of) coefficients.

[3] Though time subscripts are suppressed on all model parameters including the discount rate r, nonstochastic changes in these parameters over time is easily incorporated in the theoretical analysis and is allowed in the empirical analysis, which is conducted on annual cross-sections.

[4] Ryan (1995) provides two reasons why the accounting recognition of shocks tends to be gradual: historical cost accounting and the aggregation of accruals across vintages of assets or projects.

Substituting the expressions for BV_t and BV_{t-1} from AM into CS and substituting for $MV_t - MV_{t-1} + D_t$ in the resulting expression using NA yields the expression for earnings:

$$\begin{aligned} X_t = {} & rMV_{t-1} + (1 - F_{0p} - F_{0n} * D\phi_t)\phi_t + (F_{0p} - F_{1p} + [F_{0n} - F_{1n}]D\phi_{t-1})\phi_{t-1} \\ & + (F_{1p} - F_{2p} + [F_{1n} - F_{2n}]D\phi_{t-2})\phi_{t-2} + \cdots \qquad (1) \end{aligned}$$

This equation is easily interpreted. Earnings equal expected permanent earnings, rMV_{t-1}, plus a moving average process of current and lagged shocks to market value. The coefficient on each shock is the portion of the shock that is recognized in earnings in the current period. For example, the coefficient on the current positive (negative) shock is $1 - F_{0p}$ $(1 - F_{0p} - F_{0n})$. Moreover, the sums of the coefficients on the 1 to s period lagged positive (negative) shocks are $F_{0p} - F_{sp}$ $(F_{0p} - F_{sp} + F_{0n} - F_{sn})$, which are the portions of such shocks reflected in earnings from 1 to s periods after the shocks. These sums of coefficients are our main measures of lags.

We make two simplifying assumptions to derive our main regression equation from Equation (1). First, as in Pope and Walker (1999), we assume that $D_t = rMV_{t-1}$ (i.e., dividends equal expected permanent earnings), so that $\phi_t = MV_t - MV_{t-1}$. While this assumption serves no useful purpose for equations deflated by MV_{t-1}, it considerably simplifies the specification of equations deflated by MV_{t-4} that are estimated in Section VI. Since this assumption has no significant effect on any of our results, we maintain it throughout to enhance comparability across equations. Second, based on the previously cited research that prices lead earnings by three years, we include only the current and three lagged price changes in the regression equation. Since price changes are close to serially uncorrelated, omission of further lagged price changes should not yield correlated omitted variables problems.

Like Pope and Walker (1999), we deflate the regression equation by beginning of current period market value, MV_{t-1}. Pope and Walker (1999) also report results using MV_{t-4} as the deflator and show that the results are sensitive to the choice of deflator. MV_{t-1} is a preferable deflator for our purposes, because deflation by MV_{t-4} renders the coefficients on lagged price changes dependent on the discount rate r, as shown in Section VI, while deflation by MV_{t-1} renders only the intercept dependent on r, as can be seen directly from inspection of Equation (1). We want to interpret time-series variation in these slope coefficients as reflecting changes in lags, not discount rates. To test our hypotheses deflating by MV_{t-4}, we would have to estimate year-specific discount rates and remove them from the estimated coefficients, a process that would entail considerable error. Moreover, our main purpose is to explain results in prior studies that estimate models that deflate by MV_{t-1}. Still, to articulate with the prior literature, we also report results deflating by MV_{t-4} in Section VI.[5]

Equation (1) and these additional assumptions and specification choices yield our main regression equation, called Asym&Lags:

---

[5] Moreover, under the assumption that the undeflated model is well specified, there is no econometric reason to deflate by MV_{t-4} rather than MV_{t-1}, since the difference between these deflators is fully explained by the undeflated explanatory variables and so cannot be correlated with the undeflated residual term.

$$
\begin{aligned}
\frac{X_t}{MV_{t-1}} = {} & \alpha + \alpha_{0n}\, DR_t + \alpha_{1n}\, DR_{t-1} + \alpha_{2n}\, DR_{t-2} + \alpha_{3n}\, DR_{t-3} \\
& + \beta_{0p}\, R_t + \beta_{1p}\, R_{t-1} + \beta_{2p}\, R_{t-2} + \beta_{3p}\, R_{t-3} \qquad \text{(Asym\&Lags)} \\
& + \beta_{0n}\, (R_t * DR_t) + \beta_{1n}\, (R_{t-1} * DR_{t-1}) \\
& + \beta_{2n}\, (R_{t-2} * DR_{t-2}) + \beta_{3n}\, (R_{t-3} * DR_{t-3}) \ + \varepsilon_t,
\end{aligned}
$$

where $R_{t-s} = (MV_{t-s} - MV_{t-s-1})/MV_{t-1}$ and DR_{t-s} is a dummy variable that equals 1 if R_{t-s} is negative, and 0 otherwise. In this equation, β_{sp} and $\beta_{sp} + \beta_{sn}$ are the "total" coefficients on positive and negative returns, respectively, while β_{sn} are the "incremental" coefficients on negative returns.[6]

To simplify the empirical analysis of the Asym&Lags equation, we focus on the following (sums of) coefficients whose interpretations are provided by Equation (1). First, we interpret the estimate of β_{0p} as the portion of positive shocks that is recognized in earnings in the year the shock occurs; if there were no lags or asymmetry this would be the only significant coefficient and it would equal 1 (as would the R^2). Second, we interpret the estimate of β_{0n} as the incremental portion of negative shocks above that of positive shocks that is recognized in earnings in the year the shock occurs, i.e., as the effect of asymmetry with respect to current shocks. Third, we interpret the estimate of $\beta_{1p} + \beta_{2p} + \beta_{3p}$, denoted β_{1+2+3p}, as the portion of positive shocks that is recognized in earnings from one to three years after the shock occurs, i.e., as the effect of lags ignoring asymmetry. Fourth, we interpret the estimate of $\beta_{1n} + \beta_{2n} + \beta_{3n}$, denoted β_{1+2+3n}, as the incremental portion of negative shocks above that of positive shocks that is recognized in earnings from one to three years after the shock occurs, i.e., as the effect of asymmetry with respect to lagged shocks.

Our empirical analysis is similar to that in Pope and Walker (1999) and Ball et al. (2000), but has two important differences. First, their theoretical models do not allow the specific interpretations of the coefficients that ours does, primarily because their models do not incorporate the clean surplus relation. Second, Pope and Walker (1999) and Ball et al. (2000) are concerned with differences across countries, whereas we are interested in intertemporal changes in the U.S.[7]

To further illustrate the individual and joint effects of asymmetry and lags, we also estimate the following restricted versions of the Asym&Lags equation. As in Beaver and Ryan (1993), we estimate an equation with only lags:

$$
\frac{X_t}{MV_{t-1}} = \alpha + \beta_0\, R_t + \beta_1\, R_{t-1} + \beta_2\, R_{t-2} + \beta_3\, R_{t-3} + \varepsilon_t. \qquad \text{(Lags)}
$$

As in Basu (1997), we estimate an equation with only asymmetry:

---

[6] The Asym&Lags equation includes both intercept and slope dummies. While Equation (1) does not imply that intercept dummies are necessary, we include these to ensure that the slope coefficients do not capture mean effects.

[7] Pope and Walker (1999) report the results of annual cross-sectional estimations of one of their regression equations in their Table 6, but they do not attempt to explain intertemporal changes in the results.

$$\frac{X_t}{MV_{t-1}} = \alpha + \alpha_n DR_t + \beta_p R_t + \beta_n (R_t * DR_t) + \varepsilon_t. \qquad \text{(Asym)}$$

Similar to Beaver et al. (1987), we also estimate an equation with no asymmetry or lags:

$$\frac{X_t}{MV_{t-1}} = \alpha + \beta R_t + \varepsilon_t. \qquad \text{(Basic)}$$

In addition to analyzing (sums of) coefficients in the Asym&Lags equation, we also analyze the incremental R^2s of the less restricted equations above that of the Basic equation in order to determine the incremental and joint effects of lags and asymmetry.

All of the above regression equations are "reversed," i.e., earnings is regressed on returns rather than returns on earnings as done in most of the studies documenting the declining contemporaneous linear returns-earnings relation. While from an *information* perspective it is preferable to think of earnings as causing returns and so earnings should be the explanatory variable(s), as argued by Muller and Riedl (2001), we argue that reverse regressions provide the most straightforward way to determine the reasons for the declining association between contemporaneous returns and earnings for the following two reasons. First, we believe that this decline is much more likely to reflect changes in the nature of earnings as a summary measure of periodic firm performance than in the nature of returns, and so from this *measurement* perspective earnings is the natural dependent variable.[8] Second, prior research shows that earnings is associated with a number of lagged returns, and these close to uncorrelated returns can be included (or not) as explanatory variables in a reverse regression without substantial effect on the coefficient on any other return, yielding more interpretable coefficients than in forward regressions that allow for lags. Still, to articulate with the literature, we also estimate the more typical "forward" version of the Basic regression equation, as in Easton and Harris (1991):[9]

$$R_t = \alpha + \gamma \frac{X_t}{MV_{t-1}} + \varepsilon_t. \qquad \text{(BasicF)}$$

III. HYPOTHESES

In this section, we discuss our hypotheses regarding increasing lags and asymmetry over time. We state the hypotheses in terms of specific implications for coefficients and incremental R^2s of less restricted equations above the R^2 of the Basic equation.

Our first hypothesis is that the asymmetric response of earnings to market value shocks increases over time. As in Basu (1997), our main reason for this hypothesis is increasing accounting conservatism that reflects both the increasing number of accounting rules mandating the recognition of losses, but not gains (e.g., impairment accounting) and also managerial reporting behavior that recognizes bad news relatively quickly through big baths, in part because of legal liability for both firms and auditors. While there have been some cycles in legal liability over time, as discussed by Basu (1997), the trend is clearly up. For these (and undoubtedly other) reasons, research by Elliott and Shaw (1988), Collins et al. (1997), and Elliott and Hanna (1996), among others, shows that firms have recorded more frequent one-time items (often write-offs) over time. Asymmetry could also result from the availability of non-earnings information or market efficiency varying with the level of accounting conservatism or the presence of good and bad news (see Gigler and Hemmer 2001), or from competition yielding economic winners that take all over extended periods while losers suffer more quickly.

---

[8] Beaver (1996) discusses the roles of the information and measurement perspectives in accounting research.

[9] Beaver et al. (1987) show that Basic and BasicF yield the same t-statistics and R^2 and so convey the same information. However, $\gamma = R^2/\beta$, so that γ is reduced either when R^2 declines or β increases. While R^2 and β both reflect how much of the market value shock this period is reflected in earnings this period, they do not always move together. For example, white noise in earnings would reduce both R^2 and γ but not β.

Specifically, consistent with the findings of Basu (1997) and Givoly and Hayn (2000), we expect that the incremental coefficients on negative current returns, β_n in the Asym equation and β_{0n} in the Asym&Lags equation, become more positive over time. The sign of the sum of the incremental coefficients on negative lagged returns β_{1+2+3n} in the Asym&Lags equation is less clear, however. If more of current period bad news than of good news is recognized in the current period, then less of current period bad news than of good news must be recognized in *all* subsequent periods combined. This does not imply, however, that less current period bad news than good news is recognized in *each* subsequent period. It is entirely possible and perhaps likely that conservatism does not force the recognition of all bad news in earnings in the current year, but rather only tends to yield the recognition of bad news sooner than good news, so that β_{sn}, $s > 0$, is positive for "sufficiently short" lags and negative thereafter.[10] We also expect that the incremental R^2s of regression equations that allow for asymmetry increase over time. Basu (1997) and Givoly and Hayn (2000) do not examine incremental R^2s.

Our second hypothesis is that earnings have become increasingly untimely relative to returns. The main reasons for this hypothesis are that the historical cost valuation basis and realizability principle that underlie accounting income recognition are increasingly limited in an uncertain, competitively dynamic, and intangible-assets-intensive world, and the increasing availability and use of timelier non-earnings information for valuation purposes. Another reason for lags is that considerable anecdotal evidence and some indirect research evidence (e.g., on the increasing prevalence of one-time items mentioned above) suggests that managers' discretionary reporting behavior has increased over time. While it is logically possible that managers could apply discretion in a way that decreases lags (e.g., signals their private information), we expect that discretionary reporting behavior usually implies that earnings is less likely to reflect shocks on a timely basis and so increases lags.

Specifically, we expect that the coefficient β_{0p} on the current positive return in the Asym&Lags regression and to a lesser extent the coefficient β_0 on the current return in the Lags regression are less positive over time. We also expect that the sum of the coefficients on the positive lagged returns β_{1+2+3p} in the Asym&Lags regression and to a lesser extent the coefficients β_1, β_2, and β_3 on the lagged returns in the Lags regression are increasingly positive over time. We also expect that the incremental R^2s in regression equations with lags compared to restricted regression equations without lags increase over time.

If increasing lags and asymmetry suffice to explain the declining contemporaneous linear returns-earnings relation, then the R^2 in the Asym&Lags regression equation should be flat over time. Since prior research raises questions regarding the interpretability of a flat raw R^2 over time, we do not state this as a formal hypothesis, although we do discuss our findings regarding the trend in the R^2 in the Asym&Lags equation below because of the importance of this trend in motivating the literature.

---

[10] We make no hypotheses about how short "sufficiently short" is; Pope and Walker (1999) report significant positive coefficients on negative return slope dummies for up to one annual lag in the average of annual cross-sectional estimations and for all of the three included annual lags in the pooled estimation.

IV. SAMPLE AND DESCRIPTIVE STATISTICS

The sample includes annual data for all nonfinancial firms (i.e., firms not in SIC codes 6000–6499) that are available on the complete (i.e., including nonsurviving firms) Compustat database from 1966–2000 and that have all the variables necessary to estimate the Asym&Lags regression equation (i.e., current net income and the current and three lagged returns) in any year. These data requirements cause some survivorship bias that likely overstates the importance of lags and possibly asymmetry. Like Pope and Walker (1999), we exclude financial firms because we expect them to exhibit a significantly different association between earnings and returns than do nonfinancial firms.[11] Our sample starts in 1966 because the number of firms with the necessary data drops off significantly before then. The sample includes 105,268 observations, rising from an average of 1,483 per year in the 1966–1970 period to an average of 4,379 per year in the 1996–2000 period.

Descriptive statistics for the full sample are provided in Table 1. Panel A provides quartiles and means for X_t/MV_{t-1}, R_t, and DR_t pooled across 1966–2000. As in Pope and Walker (1999), to reduce the effects of outliers in the regression equation estimations, we delete observations in the top and bottom 1 percent of each year's distribution of deflated current earnings and each of the current and three lagged returns; our results are not substantially affected by the treatment of outliers as long as extreme observations are deleted or pulled in. Consistent with Basu (1995) and Givoly and Hayn (2000), inspection of the quartiles indicates that X_t/MV_{t-1} is skewed left, reflecting a large number of firms reporting large losses. In contrast, R_t is skewed right, likely (though not inevitably) reflecting the fact that it has a lower bound of -1. DR_t has a mean of .457 (i.e., 45.7 percent of returns are negative).

Panel B of Table 1 provides means and medians for the variables for the seven five-year periods within 1966–2000. Consistent with Givoly and Hayn (2000), inspection of differences in means and medians indicates that the left skewness of X_t/MV_{t-1} has increased dramatically over time; for example, its mean and median were both positive and approximately equal up to about 1980, but the mean has been negative since the late 1980s (e.g., $-.022$ for the 1996–2000 period), reflecting an increasingly large number of large reported losses, while the median has remained positive (e.g., .038 for the 1996–2000 period). In contrast, there is no obvious trend in the skewness of R_t; changes over time in the means and medians of R_t and DR_t appear mainly to reflect stock market conditions.

V. ANNUAL EARNINGS REGRESSION RESULTS

In this section, we test our hypotheses of increasing asymmetry and lags over time using annual earnings as the dependent variable. Table 2, Panels A–D reports annual OLS estimations of the Basic, BasicF, Asym, Lags, and Asym&Lags equations, respectively. To be able to see intertemporal trends more easily, this table reports average coefficients, t-statistics, R^2s, and number of observations for each of the seven five-year periods within

---

[11] Financial firms apply fair value accounting to a greater extent than nonfinancial firms and are subject to distinct asset impairment rules. They are disproportionately exposed to interest rate and other market risks and exhibit pervasive asset-liability matching issues. Depository institutions, the most common type of financial firm, have on average drastically reduced the duration of their (net) assets over time, which should reduce lags.

TABLE 1
Descriptive Statistics Full Sample 1966–2000

Panel A: Pooled Sample (105,268 observations from 1966–2000)

Variable	Minimum	25%	Median	75%	Maximum	Mean
X_t/MV_{t-1}	−4.637	.011	.068	.121	1.115	.031
R_t	−.986	−.218	.030	.319	12.263	.118
DR_t	0	0	0	1	1	.457

Panel B: Five-Year Periods

		Period						
Variable	Statistic	1966–1970	1971–1975	1976–1980	1981–1985	1986–1990	1991–1995	1996–2000
X_t/MV_{t-1}	mean	.063	.104	.144	.046	−.016	−.023	−.022
	median	.067	.097	.150	.084	.055	.047	.038
R_t	mean	.053	.038	.234	.152	.006	.173	.123
	median	−.029	−.040	.139	.081	−.026	.067	−.015
DR_t	mean	.534	.533	.322	.398	.510	.412	.509
	median	1	1	0	0	1	0	1
	# obs.	7415	12037	14346	15088	16351	18136	21895

Observations in the outside 1 percent of both tails of the distribution of X_t/MV_{t-1} and R_t each year are deleted.

The full sample includes all available nonfinancial firm observations (i.e., not in SIC codes 6000–6499) on Compustat from 1966–2000.

All variables necessary to estimate the Asym&Lags regression equation (see Table 2, Panel D) must be available in a year for an observation to be included in the full sample.

Variable definitions:

X_t = net income in year t;
MV_t = market value of equity at the end of fiscal year t;
R_t = fiscal year t share return; and
DR_t = a dummy variable taking a value of 1 if R_t is less than 0, and 0 otherwise.

TABLE 2
Annual OLS Estimation of Regression Equations with Annual Earnings Deflated by MV_{t-1}
Full Sample 1966–2000

Panel A: BasicF and Basic Equations

$$R_t = \alpha + \gamma \frac{X_t}{MV_{t-1}} + \varepsilon_t$$

$$\frac{X_t}{MV_{t-1}} = \alpha + \beta R_t + \varepsilon_t$$

Period	γ	β	R²	# Obs.
1996–2000	.563	.069	.05	4379
	(13.2)	(13.2)		
1991–1995	.576	.091	.05	3627
	(13.6)	(13.6)		
1986–1990	.536	.189	.10	3270
	(19.1)	(19.1)		
1981–1985	.754	.167	.13	3018
	(21.0)	(21.0)		
1976–1980	.852	.155	.13	2869
	(20.6)	(20.6)		
1971–1975	1.050	.160	.14	2407
	(19.5)	(19.5)		
1966–1970	2.816	.052	.13	1483
	(15.0)	(15.0)		
Time Trend	−.058	−.000	−.003	
	(−5.5)	(−.2)	(−4.0)	

Panel B: Asym Equation

$$\frac{X_t}{MV_{t-1}} = \alpha + \alpha_n DR_t + \beta_p R_t + \beta_n (R_t * DR_t) + \varepsilon_t$$

Period	β_p	β_n	R²	Incr. R²	# Obs.
1996–2000	−.006	.365	.12	.07	4379
	(−1.1)	(17.4)			
1991–1995	.009	.467	.12	.07	3627
	(0.9)	(14.2)			
1986–1990	.041	.476	.16	.06	3270
	(2.7)	(14.8)			
1981–1985	.063	.391	.18	.05	3018
	(5.3)	(12.0)			
1976–1980	.103	.399	.17	.04	2869
	(10.4)	(8.9)			
1971–1975	.097	.286	.17	.03	2407
	(6.3)	(7.5)			
1966–1970	.022	.069	.16	.03	1483
	(5.2)	(4.2)			
Time Trend	−.002	.010	−.001	.002	
	(−2.5)	(3.6)	(−1.8)	(4.6)	

(continued on next page)

TABLE 2 (continued)

Panel C: Lags Equation

$$\frac{X_t}{MV_{t-1}} = \alpha + \beta_0 R_t + \beta_1 R_{t-1} + \beta_2 R_{t-2} + \beta_3 R_{t-3} + \varepsilon_t$$

Period	β_0	β_1	β_2	β_3	R^2	Incr. R^2	# Obs.
1996–2000	.070	.064	.018	.005	.16	.11	4379
	(14.4)	(16.1)	(8.4)	(4.6)			
1991–1995	.105	.100	.048	.020	.20	.15	3627
	(16.7)	(17.0)	(10.8)	(6.7)			
1986–1990	.173	.109	.042	.024	.25	.15	3270
	(19.0)	(15.4)	(8.9)	(7.0)			
1981–1985	.161	.124	.063	.022	.23	.10	3018
	(21.4)	(16.2)	(8.9)	(3.4)			
1976–1980	.156	.099	.065	.045	.17	.04	2869
	(21.1)	(8.3)	(5.9)	(5.3)			
1971–1975	.174	.041	.023	.014	.20	.06	2407
	(21.3)	(6.5)	(6.2)	(3.9)			
1966–1970	.051	−.008	−.014	−.012	.19	.05	1483
	(14.9)	(−1.7)	(−3.2)	(−2.8)			
Time Trend	−.000	.002	.001	.000	.000	.003	
	(−.3)	(2.9)	(1.5)	(.9)	(.03)	(3.9)	

Panel D: Asym&Lags equation

$$\frac{X_t}{MV_{t-1}} = \alpha + \alpha_{0n} DR_t + \alpha_{1n} DR_{t-1} + \alpha_{2n} DR_{t-2} + \alpha_{3n} DR_{t-3} \div \beta_{0p} R_t + \beta_{1p} R_{t-1}$$
$$+ \beta_{2p} R_{t-2} + \beta_{3p} R_{t-3} + \beta_{0n} (R_t * DR_t) + \beta_{1n} (R_{t-1} * DR_{t-1})$$
$$+ \beta_{2n} (R_{t-2} * DR_{t-2}) + \beta_{3n} (R_{t-3} * DR_{t-3}) + \varepsilon_t$$

Period	β_{0p}	β_{0n}	β_{1+2+3p}	β_{1+2+3n}	R^2	Incr. R^2	# Obs.
1996–2000	.011	.291	.037	.033	.23	.18	4379
	(1.4)	(14.4)	(1.3)	(1.0)			
1991–1995	.042	.346	.114	.029	.25	.20	3627
	(4.6)	(11.5)	(2.6)	(1.0)			
1986–1990	.069	.339	.163	−.023	.28	.18	3270
	(4.6)	(11.0)	(3.3)	(−0.4)			
1981–1985	.073	.314	.124	.156	.28	.15	3018
	(6.6)	(10.0)	(2.7)	(3.0)			
1976–1980	.112	.370	.095	.225	.22	.09	2869
	(11.5)	(8.3)	(1.7)	(3.3)			
1971–1975	.117	.234	.052	.054	.24	.10	2407
	(8.2)	(6.5)	(0.6)	(1.4)			
1966–1970	.029	.040	−.070	.099	.22	.09	1483
	(6.4)	(2.9)	(−3.6)	(2.5)			
Time Trend	−.002	.007	.004	−.004	.001	.004	
	(−1.8)	(3.4)	(2.0)	(−1.7)	(.9)	(5.5)	

$\beta_{1+2+3p} = \beta_{1p} + \beta_{2p} + \beta_{3p}$ and $\beta_{1+2+3n} = \beta_{1n} + \beta_{2n} + \beta_{3n}$

Each regression equation is estimated using OLS each year from 1966–2000. The panels report average coefficients with average t-statistics (in parentheses), average regression equation R^2s, average incremental R^2s above the R^2 of the Basic equation, and average number of observations per year for the seven five-year periods from 1966–2000. The annual (incremental) R^2s and coefficients underlying these panels are plotted in Figures 1 and 2, respectively. The bottom rows of the panels report the coefficients with t-statistics (in parentheses) in time-series regressions of the 35 annual values of the regression equation coefficients, R^2s, and incremental R^2s on the year. See Table 1 for a description of the variables and sample.

1966–2000.[12] To test for the significance of time trends, this table also reports the results of regressions of the 35 annual (sums of) coefficients and (incremental) R^2s on the year. The t-statistics on the year in these regressions are called "time trend" t-statistics; we say that they are significant if they exceed the 5 percent level in a two-tailed test. The time trends are also depicted visually in Figure 1, which plots R^2s for all regression equations and the incremental R^2s for the Asym, Lags, and Asym&Lags equations, and Figure 2, which plots (sums of) coefficients from the Asym&Lags equation.

Panel A of Table 2 reports the estimations of the Basic and BasicF equations. These results provide a benchmark for less restricted equations and replicate previously cited prior research on the declining association between annual returns and earnings over time. The average β in the Basic equation rises from .052 in 1966–1970 to .189 in 1986–1990, unexpectedly implying that earnings reflect more of the current price change over time, before falling sharply to .069 in 1996–2000, implying the expected opposite. This U-shaped path of β yields an insignificant time trend t-statistic. As we shall see shortly, the unexpected rise in β through 1990 is primarily attributable to increasing asymmetry during this period. In contrast, the average R^2, annual values of which are plotted in Figure 1, declines almost monotonically from 13 percent in 1966–1970 to 5 percent in the 1996–2000, with a time trend $t = -4.0$, consistent with prior research. Since the coefficient γ in the forward BasicF equation equals R^2/β and β follows the U-shaped path described above, the declining R^2 is the primary reason why the average γ declines almost monotonically from 2.816 in 1966–1970 to .563 in 1996–2000, with a time trend $t = -5.5$, again consistent with prior research.

Panel B of Table 2 reports the estimation of the Asym equation. These results corroborate prior research by Basu (1997) and Givoly and Hayn (2000) regarding the importance of asymmetry with respect to the current price change. Unlike β in the Basic equation, the average coefficient β_p on R_t declines almost monotonically from .097 in 1971–1975 to $-.006$ in 1996–2000, with a significant time trend $t = -2.5$, consistent with earnings reflecting positive price changes on a less timely basis over time. Similar to Basu (1997), the average coefficient β_n on $R_t * DR_t$ is significantly positive in all of the five-year periods from 1966–2000, indicating asymmetry throughout. However, β_n rises sharply from .286 in 1971–1975 to about .47 in 1986–1995, indicating dramatically increased asymmetry during this period. While β_n decreases somewhat in the 1996–2000 period,[13] it remains at a substantially elevated level relative to the 1966–1975 period, yielding a time trend $t = 3.6$. The incremental R^2 of the Asym equation over the Basic equation, annual values of which are depicted in Figure 1, rises almost monotonically from 3 percent in 1966–1970 to 7 percent in 1996–2000, with a time trend $t = 4.6$. Thus, we find significantly increased asymmetry over time, consistent with our first hypothesis. Note that the increase in β_n drives the increase in β in the Basic equation up to 1990, since β_p declines over time.

While a partial explanation of the declining contemporaneous linear earnings-return relation, increased asymmetry is not by itself sufficient to explain the decline, since the R^2 of the Asym equation generally falls over the sample period, with a time trend $t = -1.8$.

Panel C of Table 2 reports the estimation of the Lags equation. These results corroborate prior research by Kothari and Sloan (1992), Beaver and Ryan (1993), and Collins et al. (1994), among others, regarding the importance of lags. The average coefficients β_1, β_2, and β_3 on the lagged returns are all significantly positive for each of the five-year periods

---

[12] We do not conduct any hypotheses tests based on the mean t-statistics for each five-year period; they are provided as descriptive statistics.

[13] This decrease may reflect the passage of the Private Securities Litigation Reform Act of 1995, which likely reduced the legal liability that leads to asymmetry found by Basu (1997).

FIGURE 1

Annual (Incremental R²s for Regression Equations with Annual Earnings Deflated by One-Year Lagged Price Full Sample, 1966–2000

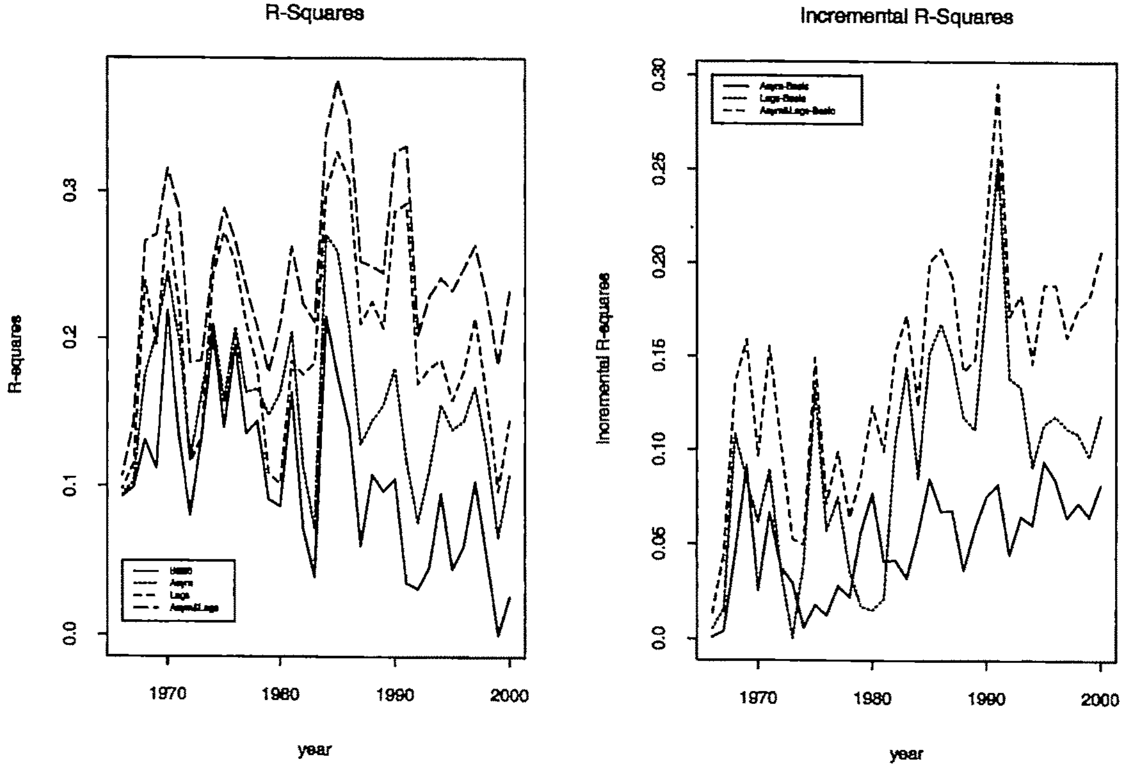
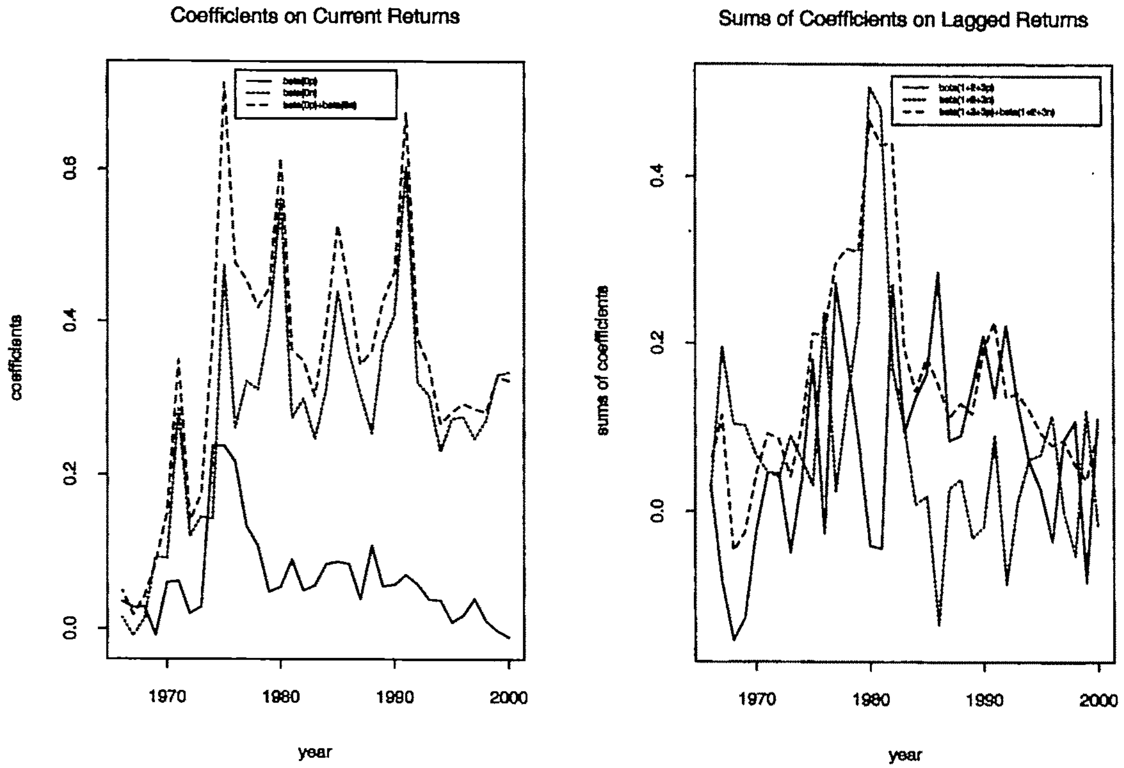


FIGURE 2
Annual Asym&Lags Equation Coefficients
Full Sample, 1966–2000



subsequent to 1970. The magnitude of the average coefficients on the first two lagged returns has generally risen over the period, with time trend t-statistics equal to 2.9 and 1.5 on β_1 and β_2, respectively, consistent with increasing lags over time. In fact, in the 1991–2000 period, β_1 is similar in magnitude to β_0, which declines sharply during the 1990s (though insignificantly over the entire period). The incremental R^2 of the Lags equation over the Basic equation, annual values of which are depicted in Figure 1, rises from 5 percent in 1966–1970 to 11 percent in 1996–2000, with a time trend $t = 3.9$. Thus, we find significantly increasing lags over time, consistent with our second hypothesis.

At first inspection, increasing lags appears to be a fairly complete explanation of the declining contemporaneous linear returns-earnings relation. In particular, the R^2 of the Lags equation does not fall over the sample period, with a time trend $t = 0.0$. However, inspection of Figure 1 indicates that this R^2 follows a U-shaped path, rising from 19 percent in 1966–1970 to 25 percent in 1986–1990 and then falling to 16 percent in 1996–2000. Moreover, previously discussed results indicate that asymmetry increases over time, something the Lags equation does not capture.

Table 2, Panel D reports the estimation of the Asym&Lags equation. Given the number of explanatory variables in this regression equation, for simplicity, only β_{0p}, β_{0n}, β_{1+2+3p}, β_{1+2+3n}, R^2, and the incremental R^2 of the Asym&Lags equation over the Basic equation are reported.

With the following exception, the trends for the coefficients in the Asym&Lags equation are consistent with those for the Asym and Lags equations, and so we do not discuss them. The average sum of the coefficients β_{1+2+3n} on the lagged negative return slope dummies is usually positive, but declines over time at the 10 percent significance level with a time trend $t = -1.7$. This implies asymmetry with respect to lagged price changes of the same direction as for the current price change, but that declines over time. This decline likely reflects the increasing asymmetry with respect to the current price change discussed previously; as more bad news is recognized in the current period, less remains to be recognized in subsequent periods.

As depicted in Figure 1, the incremental R^2 of the Asym&Lags equation over the Basic equation increases at a fast rate of .4 percent per year (time trend $t = 5.5$). There is an insignificant *increase* in the R^2 of the Asym&Lags equation over time (time trend $t = .9$), suggesting that the joint effects of asymmetry and lags largely explain the declining association between annual returns and earnings. Moreover, the R^2 exhibits considerably less of a U-shaped pattern than for the Lags equation, though some U-shape still remains with the R^2 being highest during 1981–1990. Any remaining decline in the R^2 occurs only during the 1990s.

Our finding of a flat R^2 is similar to Collins et al. (1997), who regress price on book value and earnings, and Liu and Thomas (2000), who regress returns on analyst earnings forecasts for the next five years, both of whom find that their regressions do not exhibit decreasing R^2 over time. Collins et al.'s (1997) finding is consistent with book value, as a measure of cumulative earnings plus net contributed capital, being less subject than earnings to lags, because book value has a higher signal that dwarfs the return-related "noise" caused by lags, and also asymmetry, since losses tend not to occur every year. Liu and Thomas's (2000) finding is consistent with long-term earnings forecasts being less subject than reported earnings to lags, because lags resolve over the forecast horizon, and also asymmetry, because analysts tend not to forecast one-time items, likely because of the conditional nature of asymmetry.

In summary, the results for the full sample illustrate the increasing importance of lags and asymmetry. Both effects are strong individually, but only together do they provide a reasonably complete explanation of the declining contemporaneous linear returns-earnings relation.

Constant Sample

It is possible that the declining returns-earnings relation is due, in part or whole, to changes in Compustat's coverage of firms over time. As discussed by Givoly and Hayn (2000), among others, Compustat increasingly includes younger and less profitable but higher growth potential (high-tech) firms, most notably the incorporation of NASDAQ firms in 1974. To address this possibility, we re-estimate each equation each year on a "constant" sample of firms. To be included in this sample, a firm must have all of the data necessary to estimate the Asym&Lags equation for all years from 1976 through 1999, which yields 925 firms prior to the application of deletion rules. We require continuous data starting in 1976 to allow for Compustat's incorporation of NASDAQ firms in 1974, given that Compustat typically backfills data for two years and our regression equation includes four lagged prices. If we required data availability prior to 1976, the number of firms would decline by about half. We do not require data availability in 2000 because this would eliminate some firms with non-December 31 fiscal year-ends.

While the constant sample results provide a useful check on the full sample results, they reflect two significant drawbacks and should be interpreted accordingly. First, the sample is constant only in terms of the firms' identities, not their characteristics. Second, the constant sample is far more prone to survivorship and other age-related selection biases than the full sample. These biases likely interact with both lags and asymmetry.

Since most of the distinct results for the constant sample can be observed in the Asym&Lags equation, to conserve space we report the results only for this equation (including its incremental R^2 above that of the Basic equation) in Table 3.[14] Since the sample is not constant prior to 1976, we report time trend t-statistics for both the 1966–2000 and 1976–2000 periods. The coefficients on the current returns and the incremental R^2s exhibit the same trends as for the full sample, consistent with increasing lags and asymmetry, and so we do not discuss them.

In contrast, the coefficients on the lagged returns change considerably from the full sample, perhaps reflecting selection biases induced by the construction of a constant sample. Specifically, the average sum of the coefficients on the lagged positive returns, β_{1+2+3p}, does not increase significantly over time as for the full sample, but rather follows a U-shaped path rising up to the 1976–1980 period and then falling. The average sum of the incremental coefficients on the lagged negative returns, β_{1+2+3n}, increases over both periods, albeit not quite significantly, suggesting that asymmetry with respect to lagged price changes is increasing, not decreasing as for the full sample. Not reported in Table 3, the average sum of the total coefficients on lagged negative price changes, $\beta_{1+2+3p} + \beta_{1+2+3n}$, increases significantly over the 1966–2000 period, with a time trend $t = 3.4$, consistent with increasing lags with respect to bad news.

In summary, increasing asymmetry and lags remain evident in the constant sample, suggesting that the intertemporally increasing asymmetry and decreasing timeliness that characterize the full sample are not merely due to changes in the sample composition.

---

[14] The number of firms in the constant sample results varies over time for two reasons. For all years, slightly different numbers of firms are lost due to our deletion rules. In addition, for the years up to 1975 and for 2000 firms are lost due to lack of data availability.

TABLE 3
Annual OLS Estimation of Asym&Lags Equation
Constant Sample
1966–2000

$$\frac{X_t}{MV_{t-1}} = \alpha + \alpha_{0n}\,DR_t + \alpha_{1n}\,DR_{t-1} + \alpha_{2n}\,DR_{t-2} + \alpha_{3n}\,DR_{t-3} + \beta_{0p}\,R_t + \beta_{1p}\,R_{t-1}$$

$$+ \beta_{2p}\,R_{t-2} + \beta_{3p}\,R_{t-3} + \beta_{0n}\,(R_t * DR_t) + \beta_{1n}\,(R_{t-1} * DR_{t-1})$$

$$+ \beta_{2n}\,(R_{t-2} * DR_{t-2}) + \beta_{3n}\,(R_{t-3} * DR_{t-3}) + \varepsilon_t$$

Period	β_{0p}	β_{0n}	β_{1+2+3p}	β_{1+2+3n}	R^2	Incr. R^2	# Obs.
1996–2000	.040 (3.0)	.208 (6.0)	.029 (0.8)	.133 (2.1)	.21	.16	799
1991–1995	.052 (2.7)	.191 (3.1)	.057 (0.8)	.220 (2.9)	.20	.16	812
1986–1990	.074 (4.1)	.190 (4.4)	.115 (1.9)	.036 (0.5)	.24	.10	813
1981–1985	.052 (3.6)	.195 (3.9)	.030 (0.4)	.316 (4.2)	.26	.14	808
1976–1980	.106 (8.4)	.195 (3.2)	.114 (1.3)	−.045 (−0.2)	.22	.06	808
1971–1975	.111 (4.5)	.094 (2.4)	.027 (−0.6)	.034 (1.5)	.25	.13	651
1966–1970	.027 (3.6)	.029 (1.0)	−.089 (−3.1)	.084 (1.5)	.26	.16	434
Time Trend 1966–2000	−.001 (−.7)	.005 (3.2)	.003 (1.4)	.004 (1.5)	−.001 (−1.2)	−.000 (−.2)	
Time Trend 1976–2000	−.003 (−3.0)	.001 (0.4)	−.004 (−1.2)	.006 (1.1)	−.002 (−1.3)	.003 (2.2)	

To be included in the constant sample, a firm must have all the variables required to estimated the Asym&Lags model in each year from 1976–1999. Note that the estimations cover the longer period 1966–2000.

$\beta_{1+2+3p} = \beta_{1p} + \beta_{2p} + \beta_{3p}$ and $\beta_{1+2+3n} = \beta_{1n} + \beta_{2n} + \beta_{3n}$

The Asym&Lags equation is estimated using OLS each year from 1966–2000. The table reports average coefficients with average t-statistics (in parentheses), average regression equation R^2s, average incremental R^2s above the R^2 of the Basic equation, and average number of observations per year for the seven five-year periods from 1966–2000. The bottom row reports the slope coefficients with t-statistics (in parentheses) in time-series regressions of the 35 annual values of the regression equation coefficients, R^2s, and incremental R^2s on the year.
See Table 1 for a description of the variables.

VI. FOUR-YEAR EARNINGS REGRESSION RESULTS

Our results thus far show that annual earnings exhibit increasing lags and asymmetry over time. This raises the question of whether earnings measured over longer windows exhibit increasing lags and asymmetry over time as well and, thus, whether valuation analysis using long-window earnings also increasingly needs to incorporate lags and asymmetry. Easton et al. (EHO) (1992) show that the contemporaneous linear returns-earnings relation strengthens as returns and earnings are measured over longer windows, suggesting a lesser role for lags and asymmetry over longer windows. Consistent with this, Basu (1997) shows that asymmetry is less over longer windows. However, prior research does not provide evidence on changes in lags and asymmetry for multiyear earnings over time. In this section, we provide such evidence by developing and estimating regression equations with earnings aggregated over four years as the dependent variable. We examine a four-year horizon both because it is practicable for valuation analysis, and because a longer one, such as the ten-year horizon examined by EHO, would reduce the available time series and yield survivorship bias. Basu (1997) also examines a four-year horizon.

The choice of the timing of the price deflator is critical in this analysis, since Pope and Walker (1999) find that the results of estimating regression equations similar to ours are sensitive to this choice. In this section, we deflate regressions with four-year earnings by MV_{t-4}, not by MV_{t-1} as done for the regressions with annual earnings reported in Tables 2 and 3, because we must deflate by a price no later than the beginning of the earnings window to avoid a correlation between the undeflated disturbance term and the deflator. In addition, to examine the sensitivity of our prior results to the timing of the price deflator, we re-estimate our regressions with annual earnings using MV_{t-4} as the deflator. As shown below, however, deflation by MV_{t-4} implies that the coefficients on the first three lagged returns depend on discount rates.

Extending the Theoretical Model and Regression Equations to a Four-Year Horizon

We derive our regression equations with annual and four-year earnings deflated by MV_{t-4} as the dependent variable from Equation (1). Our prior simplifying assumption that $\phi_t = MV_t - MV_{t-1}$ now becomes useful, because it implies that the term rMV_{t-1} in Equation (1) can be replaced by $r(MV_{t-s} + \phi_{t-s+1} + \phi_{t-s+2} + \cdots + \phi_{t-1})$,[15] which yields:

$$\begin{aligned} X_t = {} & rMV_{t-s} + (1 - F_{0p} - F_{0n} * D\phi_t)\phi_t + (F_{0p} - F_{1p} + [F_{0n} - F_{1n}]D\phi_{t-1} + r)\phi_{t-1} \\ & + (F_{1p} - F_{2p} + [F_{1n} - F_{2n}]D\phi_{t-2} + r)\phi_{t-2} + \cdots \\ & + (F_{s-2p} - F_{s-1p} + [F_{s-2n} - F_{s-1n}]D\phi_{t-s+1} + r)\phi_{t-s+1} \qquad (1') \\ & + (F_{s-1p} - F_{sp} + [F_{s-1n} - F_{sn}]D\phi_{t-s})\phi_{t-s} \\ & + (F_{sp} - F_{s+1p} + [F_{sn} - F_{s+1n}]D\phi_{t-s-1})\phi_{t-s-1} + \cdots \end{aligned}$$

Note that the coefficients on the shocks from 1 to $s-1$ lags are *r plus* the corresponding coefficients in Equation (1), because these shocks, reinvested at the normal rate of return r, are expected to yield future permanent earnings that are not reflected in rMV_{t-s}.

The dependence of slope coefficients on the discount rate makes it difficult to interpret time-series variation in those coefficients in terms of changes in lags and asymmetry, as we did in the prior analysis.[16] Discount rates vary substantially over our sample period, being much higher during the mid-1970s to early 1980s than subsequently, and so the coefficients on lagged returns in the MV_{t-4} equations tend to decrease in the latter half of our sample period, working in the opposite direction of our hypotheses regarding increasing lags. Prior researchers who deflate by price prior to $t-1$ have not appreciated the dependence of the coefficients on the discount rate.

---

[15] If we did not assume that $\phi_t = MV_t - MV_{t-1}$, then Equation (1′) would include the additional terms $D_{t-s} - rMV_{t-s-1}$, $1 \leq s \leq 3$, since rMV_{t-1} would be replaced by $r(MV_{t-s} + \Sigma_{s=1}^{3} [\phi_{t-s} - D_{t-s} + rMV_{t-s-1}])$ in the derivation of that equation.

[16] In Tables 2 and 3, we report the estimation of regression equations deflated by MV_{t-1}, for which only the intercepts should depend on discount rates. In fact, the intercepts in these regressions (not reported in those tables) do rise and fall along with discount rates.

We derive the unrestricted regression equation with one-year earnings directly from Equation (1′) setting $s = 4$, deflating by MV_{t-4}, and redefining R_{t-s} as ϕ_{t-s}/MV_{t-4}:

$$\begin{aligned}\frac{X_t}{MV_{t-4}} &= \alpha + \alpha_{0n}\,DR_t + \alpha_{1n}\,DR_{t-1} + \alpha_{2n}\,DR_{t-2} + \alpha_{3n}\,DR_{t-3}\\ &\quad + \beta_{0p}\,R_t + \beta_{1p}\,R_{t-1} + \beta_{2p}\,R_{t-2} + \beta_{3p}\,R_{t-3} \qquad \text{(Asym\&Lags}'\text{)}\\ &\quad + \beta_{0n}\,(R_t * DR_t) + \beta_{1n}\,(R_{t-1} * DR_{t-1})\\ &\quad + \beta_{2n}\,(R_{t-2} * DR_{t-2}) + \beta_{3n}\,(R_{t-3} * DR_{t-3}) + \varepsilon_t\end{aligned}$$

The restricted versions of this equation, denoted Basic′, Asym′, and Lags′, are analogous to the Basic, Asym, and Lags equations aside from being deflated by MV_{t-4}, and so to conserve space are not reproduced here.

The theoretical model with four-year earnings, $X_t + X_{t-1} + X_{t-2} + X_{t-3}$, in the numerator of the dependent variable is obtained by adding Equation (1′) across these four years' earnings, setting $s = 4$ on the right-hand side of the equation regardless of which year's earnings is on the left-hand side of the equation:

$$\begin{aligned}\sum_{s=0}^{3} X_{t-s} &= 4rMV_{t-4} + (1 - F_{0p} - F_{0n} * D\phi_t)\phi_t + (1 - F_{1p} - F_{1n} * D\phi_{t-1} + r)\phi_{t-1}\\ &\quad + (1 - F_{2p} - F_{2n} * D\phi_{t-2} + 2r)\phi_{t-2} + (1 - F_{3p} - F_{3n} * D\phi_{t-3} + 3r)\phi_{t-3}\\ &\quad + (F_{0p} - F_{4p} + [F_{0n} - F_{4n}]D\phi_{t-4})\phi_{t-4} \qquad (2)\\ &\quad + (F_{1p} - F_{5p} + [F_{1n} - F_{5n}]D\phi_{t-5})\phi_{t-5} + \cdots\end{aligned}$$

The coefficients in Equation (2) are interpreted easily. The coefficient on MV_{t-4} is $4r$ because the normal return on this investment is reflected in four years' earnings in the dependent variable. The coefficient on the current shock ϕ_t is the portion of that shock reflected in X_t (e.g., $1 - F_{0p}$ for positive shocks), exactly as in Equations (1) and (1′). The coefficients on the first three lagged shocks ϕ_{t-s}, $1 \leq s \leq 3$, is the portion of those shocks reflected in earnings in years $t-s+1$ to t (e.g., $1 - F_{sp}$ for positive shocks) plus the expected permanent earnings in years $t-s+1$ to t attributable to reinvestment of the shock at the normal rate of return r (i.e., sr).[17] The coefficients on further lagged shocks ϕ_{t-s}, $s \geq 3$, reflect the portion of those shocks reflected in earnings in years $t-3$ to t (e.g., $F_{s-4p} - F_{sp}$ for positive shocks).

We emphasize that the aggregation of earnings across longer windows makes our assumption of deterministic lags in AM less restrictive, since the timing of the recognition of a shock within a longer earnings window is inherently more predictable (i.e., less stochastic or subject to managerial choice).[18] This works to raise coefficients and R^2s in the regressions with four-year earnings compared to the regressions with annual earnings.

Equation (2) implies that annual returns should not be aggregated into a single aggregate return on the right-hand side, as in EHO, because each of the annual returns has a distinct association with aggregate earnings on the left-hand side.[19] Accordingly, we include annual returns in regression equations as separate independent variables in the equations below. The unrestricted regression equation with four-year earnings is derived by deflating Equation (2) by MV_{t-4} and including six lagged annual returns, so that the first year's earnings has three lagged annual returns preceding it, as in the Asym&Lags equation:

---

[17] The linear increase in the coefficients on the lagged shocks attributable to their effect on expected future permanent earnings as the lag increase from 1 to 3 (i.e., sr), is characteristic of the temporal aggregation of permanent earnings, as shown in Ryan (1988). Intuitively, shocks earlier in the earnings measurement window have a larger effect on earnings than do shocks later in the window.

[18] Stochastic lags could result from the recognition of earnings at the completion of a project of uncertain length. Managed lags could result from the recognition of earnings at events (e.g., the decision to sell appreciated securities) or using estimates determined by management.

$$\begin{aligned}\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = {} & \alpha + \alpha_{0n}\, DR_t + \alpha_{1n}\, DR_{t-1} + \alpha_{2n}\, DR_{t-2} + \cdots + \alpha_{6n}\, DR_{t-6} \\ & + \beta_{0p}\, R_t + \beta_{1p}\, R_{t-1} + \beta_{2p}\, R_{t-2} + \cdots + \beta_{6p}\, R_{t-6} \\ & + \beta_{0n}\, (R_t * DR_t) + \beta_{1n}\, (R_{t-1} * DR_{t-1}) \\ & + \beta_{2n}\, (R_{t-2} * DR_{t-2}) + \cdots + \beta_{6n}\, (R_{t-6} * DR_{t-6}) + \varepsilon_t.\end{aligned} \tag{Asym\&Lags4}$$

We specify the restricted versions of this equation as follows. The model with neither lags nor asymmetry includes the four annual returns within the earnings window:

$$\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = \alpha + \beta_0\, R_t + \beta_1\, R_{t-1} + \beta_2\, R_{t-2} + \beta_3\, R_{t-3} + \varepsilon_t. \tag{Basic4}$$

While the Basic4 equation is consistent with Equation (2), in order to maintain comparability with prior research by EHO, we also estimate a restricted version of this equation with the four annual returns aggregated into one four-year return as the independent variable defined as $R_{t,t-3} = (MV_t - MV_{t-4})/MV_{t-4}$:

$$\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = \alpha + \beta\, R_{t,t-3} + \varepsilon_t. \tag{BasicEHO}$$

We specify the equation with only asymmetry as the Basic4 equation adding negative return intercept and slope dummies for each of the returns:

$$\begin{aligned}\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = {} & \alpha + \alpha_{0n}\, DR_t + \alpha_{1n}\, DR_{t-1} + \alpha_{2n}\, DR_{t-2} + \alpha_{3n}\, DR_{t-3} \\ & + \beta_{0p}\, R_t + \beta_{1p}\, R_{t-1} + \beta_{2p}\, R_{t-2} + \beta_{3p}\, R_{t-3} \\ & + \beta_{0n}\, (R_t * DR_t) + \beta_{1n}\, (R_{t-1} * DR_{t-1}) \\ & + \beta_{2n}\, (R_{t-2} * DR_{t-2}) + \beta_{3n}\, (R_{t-3} * DR_{t-3}) + \varepsilon_t.\end{aligned} \tag{Asym4}$$

We specify the equation with only lags with the current and six lagged annual returns:

---

[19] This point applies to earnings measured over any period, even a year or a quarter, though its importance rises with the length of the period over which earnings is measured.

$$\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = \alpha + \beta_0 R_t + \beta_1 R_{t-1} + \beta_2 R_{t-2} + \beta_3 R_{t-3} + \cdots + \beta_6 R_{t-6} + \varepsilon_t. \qquad \text{(Lags4)}$$

Empirical Analysis

To ensure that differences across equations are not attributable to different samples, we require that all variables necessary to estimate the Asym&Lags4 equation be available to estimate any of the equations, deleting observations in the outside 1 percent of the distribution of any variable in any of the MV_{t-4} deflated equations. Since the Asym&Lags4 equation includes three more lags of both earnings and returns than the Asym&Lags equation, these estimations include about two-thirds the observations in the regression results reported in Table 2.

Since the coefficients on the first three lagged returns in the MV_{t-4} deflated equations depend on discount rates, we focus our analysis on the (incremental) R^2s, with the four-year and annual earnings regressions reported in Tables 4 and 5, respectively.[20] Specifically, Panel A of Table 4 reports the R^2s of the BasicEHO, Basic4, Asym4, Lags4, and Asym&Lags4 equations, while Panel B reports the incremental R^2s of the Asym4, Lags4, and Asym&Lags4 equations above that of the Basic4 equation. Similarly, Panel A of Table 5 reports the R^2s of the Basic′, Asym′, Lags′, and Asym&Lags′ equations, while Panel B reports the incremental R^2s of the Asym', Lags', and Asym&Lags' equations above that of the Basic' equation. As before, time trend t-statistics indicate the significance of changes in (incremental) R^2s over time. The (incremental) R^2s reported in Tables 4 and 5 are plotted in Figure 3 (Figure 4).[21]

Panel A of Table 4 shows that the R^2 averages about .08 higher for the Basic4 equation than for the BasicEHO equation, illustrating the importance of allowing the coefficients on the annual returns within the earnings window to vary. Most important, while the R^2s for both the Basic4 and BasicEHO equations exhibit a U-shaped pattern, these R^2s decline significantly over time (time trend t's equal −2.6 for BasicEHO and −3.5 for Basic4). Inspection of Figure 3 indicates that until the mid-1980s the R^2s of the BasicEHO and Basic4 equations are fairly flat, averaging about .32 and .40, respectively, and they fall sharply to much lower average levels of .15 and .23, respectively, in the 1996–2000 period. Thus, aggregation of earnings over a four-year window solves less of the declining contemporaneous linear returns-earnings relation over time.

Panel B of Table 4 shows that the incremental R^2s of the all of the equations rise significantly over time (time trend t-statistics equal 6.6 for Asym4, 4.0 for Lags4, and 6.7 for Asym&Lags4). Inspection of Figures 3 and 4 again indicates that the incremental R^2 of the Asym&Lags4 equation is fairly flat through the mid-1980s, after which time it rises sharply, with more of the rise attributable to asymmetry than lags. These results imply that Basu's (1997) finding that aggregation of earnings reduces asymmetry is weakening over time, consistent with bad news being recognized quicker than good news but not necessarily in the same period as the bad news occurs. The incremental R^2 for the Asym&Lags4 equation rises .4 percent per year on average, which is large enough that the R^2 in this equation does not decline significantly over time, as reflected in a time trend $t = -1.3$ in Panel A of Table 4.

---

[20] These R^2s also may reflect discount rates, however, because the permanent portion of earnings associated with the reinvestment of lagged shocks should be relatively easy to explain, and will be larger in absolute value when discount rates are higher, all else being equal.

[21] While not reported in the tables, we also estimated analogous regression equations with earnings aggregated over two-year and three-year windows. As expected, the (incremental) R^2s for earnings aggregated at these windows change smoothly as the return window lengthens.

The MV_{t-4} deflated regressions with annual earnings are reported in Table 5. Not surprisingly, the R^2s for these equations in Panel A are lower than the R^2s for the corresponding equations for four-year earnings reported in Table 4, reflecting the fact that it is easier to explain longer-window earnings. More interestingly, these results exhibit two significant differences from those for the MV_{t-1} deflated equations with annual earnings reported in Table 2, consistent with the findings of Pope and Walker (1999). First, the R^2 for the Basic′ equation in Panel A (and for all equations in this table) does not decline significantly over time, unlike the R^2 for the Basic equation in Table 2. In particular, very high R^2s obtain for the Basic′ equation in and around the high discount rate period from 1976–1980, so that the R^2s follow a U-shaped path. We attribute this difference primarily to the dependence of the coefficients in MV_{t-4} deflated equations on discount rates.

Second, the incremental R^2s in the equations involving lags (and to a lesser extent asymmetry) are uniformly very high from 1966–1985 and then, unlike the results for the multiyear earnings in Table 4, decline considerably, so that the incremental R^2s all decline

TABLE 4

Annual OLS Estimation of Regression Equations with Four-Year Earnings Deflated by MV_{t-4}

Full Sample 1966–2000

Panel A: R^2s

Period	**BasicEHO**	**Basic4**	**Asym4**	**Lags4**	**Asym&Lags4**	**# Obs.**
1996–2000	.15	.23	.35	.29	.39	2833
1991–1995	.19	.24	.34	.33	.41	2692
1986–1990	.33	.39	.47	.45	.51	2305
1981–1985	.34	.38	.45	.39	.47	2265
1976–1980	.38	.48	.51	.50	.53	2261
1971–1975	.26	.35	.41	.38	.45	1640
1966–1970	.27	.35	.39	.37	.41	920
Time Trend	−.004	−.005	−.003	−.003	−.002	
	(−2.6)	(−3.5)	(−2.0)	(−2.3)	(−1.3)	

Panel B: Incremental R^2s (relative to Basic4)

Period	**Asym4**	**Lags4**	**Asym&Lags4**
1996–2000	.12	.06	.17
1991–1995	.10	.10	.17
1986–1990	.08	.06	.12
1981–1985	.07	.01	.09
1976–1980	.03	.01	.04
1971–1975	.06	.03	.10
1966–1970	.04	.02	.07
Time Trend	.003	.002	.004
	(6.6)	(4.0)	(6.7)

(*continued on next page*)

TABLE 4 (continued)

The sample includes all firm-year observations from 1966–2000 with the data necessary to estimate the Asym&Lags4 equation, deleting observations in the outside 1 percent of the distributions of X_t/MV_{t-4}, $\Sigma_{s=0}^{3} X_{t-s}/MV_{t-4}$, R_t each year, and $R_{t,t-3}$ to ensure comparability of the results in Tables 4 and 5.
The regression equations are (where $R_{t,t-3}$ denotes $(MV_t - MV_{t-4})/MV_{t-4}$):

$$\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = \alpha + \beta\, R_{t,t-3} + \varepsilon_t. \tag{BasicEHO}$$

$$\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = \alpha + \beta_0\, R_t + \beta_1\, R_{t-1} + \beta_2\, R_{t-2} + \beta_3\, R_{t-3} + \varepsilon_t \tag{Basic4}$$

$$\begin{aligned}\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = {} & \alpha + \alpha_{0n}\, DR_t + \alpha_{1n}\, DR_{t-1} + \alpha_{2n}\, DR_{t-2} + \alpha_{3n}\, DR_{t-3} \\ & + \beta_{0p}\, R_t + \beta_{1p}\, R_{t-1} + \beta_{2p}\, R_{t-2} + \beta_{3p}\, R_{t-3} \\ & + \beta_{0n}\, (R_t * DR_t) + \beta_{1n}\, (R_{t-1} * DR_{t-1}) \\ & + \beta_{2n}\, (R_{t-2} * DR_{t-2}) + \beta_{3n}\, (R_{t-3} * DR_{t-3}) + \varepsilon_t \end{aligned} \tag{Asym4}$$

$$\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = \alpha + \beta_0\, R_t + \beta_1\, R_{t-1} + \beta_2\, R_{t-2} + \beta_3\, R_{t-3} + \cdots + \beta_6\, R_{t-6} + \varepsilon_t \tag{Lags4}$$

$$\begin{aligned}\frac{\sum_{s=0}^{3} X_{t-s}}{MV_{t-4}} = {} & \alpha + \alpha_{0n}\, DR_t + \alpha_{1n}\, DR_{t-1} + \alpha_2 DR_{t-2} + \cdots + \alpha_{6n}\, DR_{t-6} \\ & + \beta_{0p}\, R_t + \beta_{1p}\, R_{t-1} + \beta_{2p}\, R_{t-2} + \cdots + \beta_{6p}\, R_{t-6} \\ & + \beta_{0n}\, (R_t * DR_t) + \beta_{1n}\, (R_{t-1} * DR_{t-1}) \\ & + \beta_{2n}\, (R_{t-2} * DR_{t-2}) + \cdots + \beta_{6n}\, (R_{t-6} * DR_{t-6}) + \varepsilon_t. \end{aligned} \tag{Asym\&Lags4}$$

Each equation is estimated using OLS each year from 1966–2000. The panels report average R^2s and incremental R^2s above the R^2 of the Basic4 equation and average number of observations for the seven five-year periods from 1966–2000. The R^2s and incremental R^2s reported in these panels are plotted in Figures 3 and 4, respectively. The bottom rows of the panels report the slope coefficients with t-statistics (in parentheses) in time-series regressions of the 35 annual values of the R^2s and incremental R^2s on the year.

significantly over time (time trend t's equal −1.9 for Asym′, −3.3 for Lags′, and −2.5 for Asym&Lags′). While this likely reflects trends in the path of discount rates to some extent, discount rates cannot explain the high incremental R^2s in 1966–1970 period, when discount rates were low. For some reason, it is much easier to explain X_t/MV_{t-4} than X_t/MV_{t-1} early in the sample period. We conjecture that the explanation is related to the fact that to the extent that X_t is permanent, X_t/MV_{t-4} has characteristics of long-window earnings (see Ryan and Zarowin 1995), and that X_t is more permanent, and so X_t/MV_{t-4} is more explainable, early in the sample period. More research is clearly needed to reconcile the results for the MV_{t-4} and MV_{t-1} deflated equations. The results in Table 5 do show, however, that the R^2 in the Basic′ equation declines sharply from 1976–1980 period on, and that incorporating asymmetry and especially lags into the equation contribute considerable incremental R^2.

TABLE 5
Annual OLS Estimation of Regression Equations with Annual Earnings Deflated by MV_{t-4}
Full Sample 1966–2000

Panel A: R^2s

Period	Basic′	Asym′	Lags′	Asym&Lags′	# Obs.
1996–2000	.07	.09	.17	.25	2833
1991–1995	.11	.13	.24	.30	2692
1986–1990	.13	.14	.30	.35	2305
1981–1985	.14	.18	.35	.39	2265
1976–1980	.21	.23	.43	.46	2261
1971–1975	.10	.13	.30	.35	1640
1966–1970	.12	.18	.35	.38	920
Time Trend	−.002 (−1.1)	−.003 (−2.0)	−.006 (−3.8)	−.004 (−3.5)	

Panel B: Incremental R^2s (relative to Basic′)

Period	Asym′	Lags′	Asym&Lags′
1996–2000	.02	.10	.18
1991–1995	.02	.13	.19
1986–1990	.01	.17	.22
1981–1985	.04	.21	.25
1976–1980	.02	.22	.25
1971–1975	.03	.20	.25
1966–1970	.05	.23	.26
Time Trend	−.001 (−1.9)	−.004 (−3.3)	−.003 (−2.5)

The regression equations are the same as those estimated in Table 2 except that the deflator for all variables is MV_{t-4} not MV_{t-1}. Primes are placed at the end of the equation names to indicate the different deflator.
Each equation is estimated using OLS each year from 1966–2000. The panels report average R^2s and incremental R^2s above the R^2 of the Basic′ equation and average number of observations for the seven five-year periods from 1966–2000. The R^2s and incremental R^2s reported in these panels are plotted in Figures 3 and 4, respectively. The bottom rows of the panels report the slope coefficients with t-statistics (in parentheses) in time-series regressions of the 35 annual values of the R^2s and incremental R^2s on the year.
The sample is described in Table 4.

In summary, the results in this section imply that, since the mid-1980s, aggregation of earnings across a medium-long window of four years solves increasingly less of the low and declining contemporaneous linear returns-earnings relation.

VII. ANNUAL CASH FLOW FROM OPERATIONS REGRESSION RESULTS

To test whether our results on the increasing importance of lags and asymmetry are attributable to accounting reasons reflected in accruals vs. economic reasons reflected in cash flow from operations (CFO), in this section we re-estimate our annual MV_{t-1} deflated regression equations using annual CFO as the dependent variable, as in Basu (1997). These results begin in 1987, when firms were first required to report CFO under SFAS No. 87, because the estimation of CFO prior to 1987 requires first estimating accruals using the "balance sheet method" and calculating CFO as income (or funds flow) minus (current) accruals. Collins and Hribar (2002) show that this method can cause large errors in the

FIGURE 3
Annual R²s for Regression Equations with Annual and Four-Year Earnings Deflated by Four-Year Lagged Price
Full Sample, 1966–2000

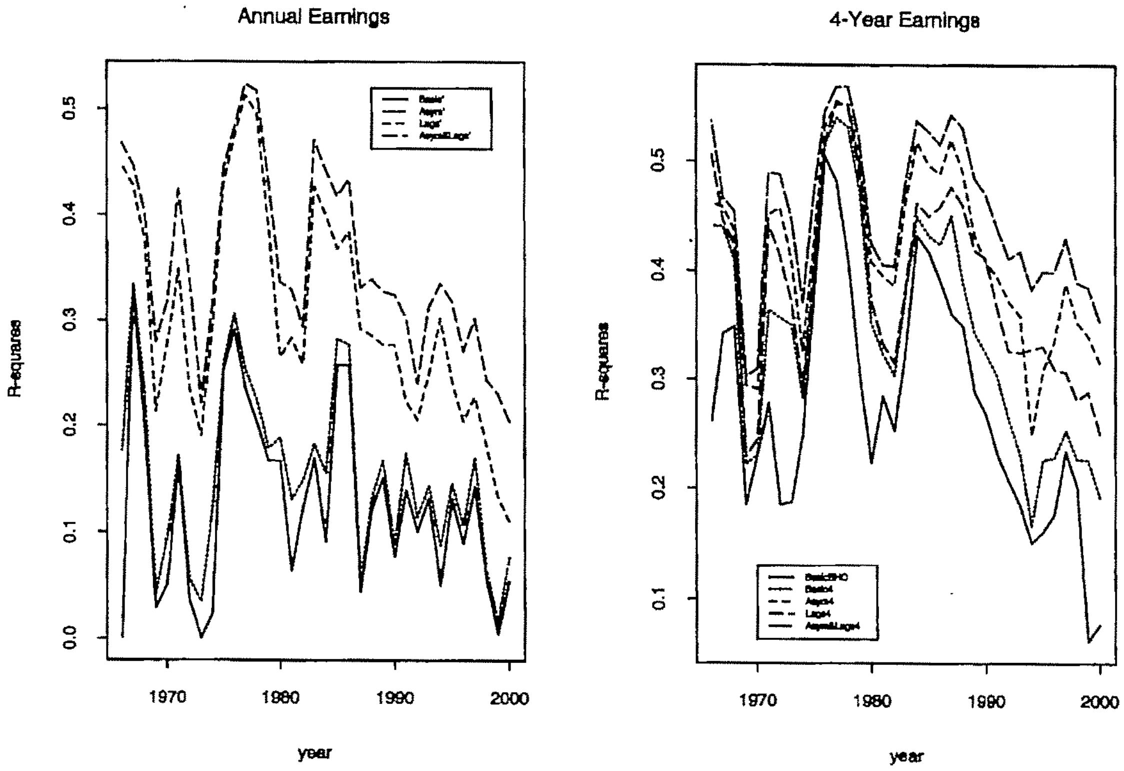
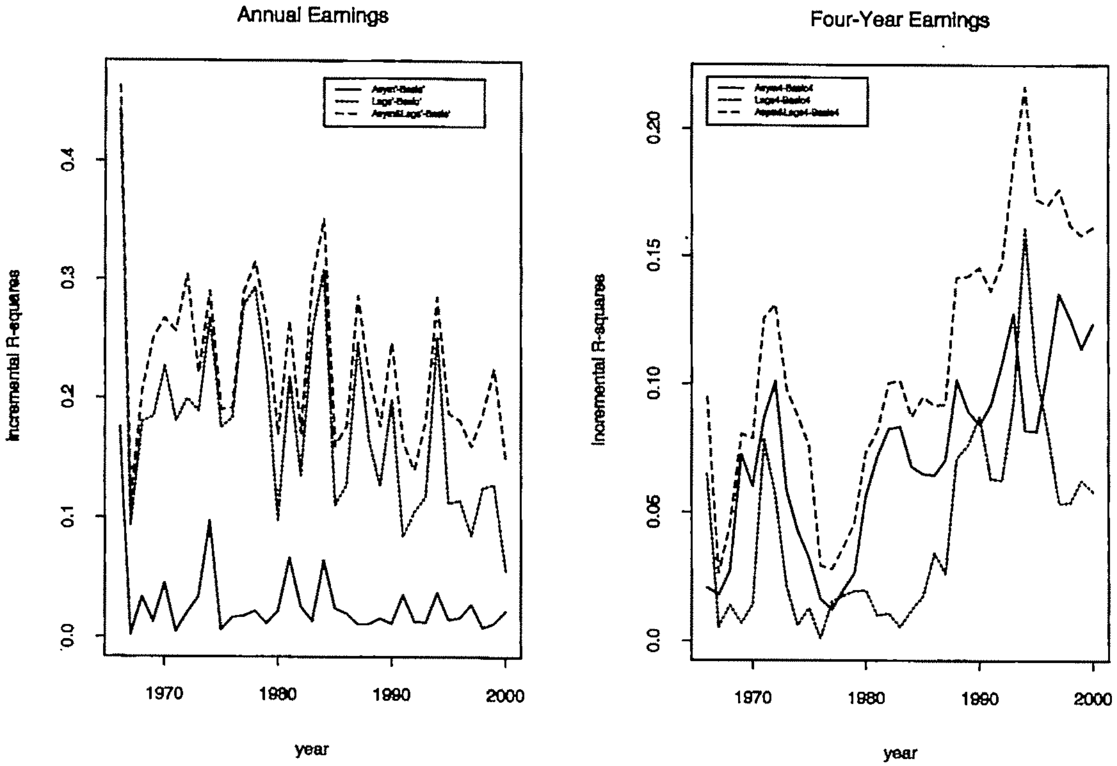


FIGURE 4
Annual Incremental R²s for Regression Equations with Annual and Four-Year Earnings Deflated by Four-Year Lagged Price
Full Sample, 1966–2000



estimates of accruals for firms engaging in significant mergers and acquisitions and discontinuance of operations.[22]

Since all of the effects of substituting CFO for earnings are observable in the results of estimating the Asym&Lags equation, Table 6 reports only the results of estimating this equation on the full sample. For comparability, the regression results are averaged over the same five-year periods as in Table 2, with only four years in the first period 1987–1990.

The coefficient β_{0p} on the current positive price change is positive, similar in magnitude to that coefficient in Panel D of Table 2, and similarly declining over time (time trend t = −3.8), implying that, like earnings, CFO exhibits more lags over time. The coefficient β_{0n} on the current negative return slope dummy is positive but considerably lower in magnitude (about one-quarter in 1987–1995 and less than one-half in 1996–2000) than in Panel D of Table 2, implying that CFO exhibits asymmetry with respect to the current price change but at a much lower level than earnings. This coefficient rises over time (time trend t = 2.6), implying that, like earnings, CFO exhibits increasing asymmetry over time.

TABLE 6
Annual OLS Estimation of Asym&Lags Equation with Annual Cash Flow from Operations Deflated by MV_{t-1}
Full Sample
1987–2000

$$\frac{CFO_t}{MV_{t-1}} = \alpha + \alpha_{0n}\, DR_t + \alpha_{1n}\, DR_{t-1} + \alpha_{2n}\, DR_{t-2} + \alpha_{3n}\, DR_{t-3} + \beta_{0p}\, R_t + \beta_{1p}\, R_{t-1} + \beta_{2p}\, R_{t-2} + \beta_{3p}\, R_{t-3} + \beta_{0n}\, (R_t * DR_t) + \beta_{1n}\, (R_{t-1} * DR_{t-1}) + \beta_{2n}\, (R_{t-2} * DR_{t-2}) + \beta_{3n}\, (R_{t-3} * DR_{t-3}) + \varepsilon_t$$

Period	β_{0p}	β_{0n}	β_{1+2+3p}	β_{1+2+3n}	R^2	Incr. R^2	# Obs.
1996–2000	.007 (0.7)	.135 (7.0)	−.147 (−4.9)	.170 (5.8)	.09	.07	4379
1991–1995	.022 (2.2)	.092 (3.4)	−.181 (−5.0)	.188 (5.4)	.06	.04	3627
1987–1990	.066 (3.2)	.090 (2.8)	−.127 (−2.2)	.144 (2.6)	.07	.03	3270
Time Trend	−.006 (−3.8)	.004 (2.6)	.000 (0.1)	.000 (0.1)	.001 (.9)	.004 (4.5)	

$\beta_{1+2+3p} = \beta_{1p} + \beta_{2p} + \beta_{3p}$ and $\beta_{1+2+3n} = \beta_{1n} + \beta_{2n} + \beta_{3n}$

The Asym&Lags equation is estimated with cash flow from operations as reported on the cash flow statement as the dependent variable using OLS each year from 1987–2000. The table reports average coefficients with average t-statistics (in parentheses), average regression equation R^2s, average incremental R^2s above the R^2 of the Basic equation, and average number of observations per year for the periods 1987–1990, 1991–1995, and 1996–2000. The bottom row reports the slope coefficients with t-statistics (in parentheses) in time-series regressions of the 14 annual values of the regression equation coefficients, R^2s, and incremental R^2s on the year.
See Table 1 for a description of the variables and sample.

---

[22] We estimated the Asym&Lags equation over 1987–2000 using the deflated *difference* between CFO estimated by the balance sheet method and CFO from the cash flow statement (i.e., the *error* in CFO estimated by the balance sheet method) as the dependent variable. We found that this error exhibits both lags and asymmetry more strongly than does CFO from the cash flow statement. Moreover, there are significant trends in this error's lags and asymmetry over time.

The average sum of the coefficients β_{1+2+3_p} on the lagged positive returns is consistently significantly negative, implying dramatically different lags for CFO than for earnings. While this result may seem surprising, it may reflect successful firms choosing to grow in ways that reduce their CFO. There is no significant trend in β_{1+2+3_p} over time. The average sum of the coefficients β_{1+2+3_n} on the lagged negative returns is consistently significantly positive, implying that like earnings CFO reflects lagged news asymmetrically. However, note that β_{1+2+3_p} and β_{1+2+3_n} are approximately equal in absolute magnitude but have opposite sign in each period, so that lagged negative shocks have essentially no implications for CFO.

Both the R^2 and incremental R^2 in the Asym&Lags equation are considerably smaller in Table 6 than in Panel D of Table 2. For example, the incremental R^2 in the 1996–2000 period is .07 compared to .18 in Panel D of Table 2. Lags and asymmetry each contribute about half of this incremental R^2. Moreover, this smaller incremental R^2 obtains despite a much smaller R^2 in the Basic equation in 1996–2000 of .02 compared to .05 in Panel D of Table 2. The incremental R^2 in the Asym&Lags equation does rise over time; however, with a time trend $t = 4.5$, implying that CFO exhibits increasing lags and asymmetry.

In summary, while CFO exhibits lags and asymmetry that increase over time, much less of CFO than earnings is explained by lags and asymmetry. Moreover, CFO and earnings' lags are of a very different nature. Thus earnings' lags and asymmetry reported in Table 2 primarily reflect accounting reasons reflected in accruals rather than economic reasons reflected in CFO.[23]

VIII. CONCLUSION

In this paper, we find that the declining contemporaneous linear relation between annual stock returns and earnings documented by prior research is largely explained by increases in earnings' lags and asymmetry over time. We find that the increasing importance of lags and asymmetry remains when earnings is aggregated over a four-year period, and that as a consequence the relation between four-year earnings and returns is very weak by the end of our sample period. We find that lags (albeit of a different sort) and asymmetry also obtain on a constant sample of firms, and so are not solely attributable to changes in sample composition. We find stronger lags and asymmetry for the accrual component than the cash flow from operations component of earnings, implying our results are primarily attributable to accounting, not economic, reasons. We also make two related methodological contributions, first by developing a theoretical model that allows us to assess lags and asymmetry through examination of coefficients (not R^2s) in regressions of earnings on current and lagged annual returns, and second by demonstrating the dependence of the slope coefficients on discount rates when lagged price deflators are used.

Our results should help accounting policymakers and users of financial reports identify the reasons for the declining returns-earnings relation over time, and thus allow them to decide what, if anything, should be done to address the situation. For example, accounting standards setters might find lags more disturbing than asymmetry, given the importance they attach to the conservatism principle. In contrast, users of financial reports might expect or cause lags, given their awareness and use of a wide range of non-earnings information. In another sense, however, our results have simply replaced one question—Why has the contemporaneous linear returns-earnings relation declined?—with two more refined questions—Why have each of lags and asymmetry increased? Both lags and asymmetry could result from the nature of accounting rules, managerial reporting choices, non-earnings information, market inefficiency, or competition.

---

[23] Our results that CFO exhibits lags and asymmetry like earnings are consistent with Kim and Kross's (2002) finding that the association between earnings and future CFO strengthens over time.

For example, increasing lags could reflect the increasing limitations of the historical cost valuation basis and of the realizability principle that underlie accounting income recognition in an uncertain, competitively dynamic, and intangible-assets-driven world, the increasing availability and use of timelier non-earnings information for valuation purposes, or increases in certain types of managerial discretionary reporting behavior. Similarly, increasing asymmetry could reflect increases in the conservatism of accounting rules or in managerial reporting behavior given those rules and the presence of legal liability, the availability of non-earnings information or market efficiency varying with the level of accounting conservatism or the presence of good and bad news, or competition yielding economic winners that take all over extended periods, while losers suffer more quickly. Future research is needed to assess how these factors, individually and collectively, answer these two more refined questions.

REFERENCES

Ball, R., S. P. Kothari, and A. Robin. 2000. The effect of international institutional factors on properties of accounting earnings. *Journal of Accounting and Economics* 29: 1–51.

Basu, S. 1995. Conservatism and the asymmetric timeliness of earnings. Ph.D. dissertation, University of Rochester.

———. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24: 3–37.

Beaver, W., R. Lambert, and S. Ryan. 1987. The information content of security prices: A second look. *Journal of Accounting and Economics* 9: 139–157.

———, and S. Ryan. 1993. Accounting fundamentals of the book-to-market ratio. *Financial Analysts Journal* (November/December): 50–56.

———. 1996. Directions in accounting research: NEAR and FAR. *Accounting Horizons* 10 (June): 113–124.

Brown, S., K. Lo, and T. Lys. 1999. Use of R^2 in accounting research: Measuring changes in value relevance over the last four decades. *Journal of Accounting and Economics* 28: 83–115.

Collins, D., S. P. Kothari, J. Shanken, and R. Sloan. 1994. Lack of timeliness and noise as explanations for the low contemporaneous return-earnings association. *Journal of Accounting and Economics* 18: 289–324.

———, E. Maydew, and I. Weiss. 1997. Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics* 27: 39–67.

———, and P. Hribar. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (March): 105–134.

Easton, P., and T. Harris. 1991. Earnings as an explanatory variable for returns. *Journal of Accounting Research* 29: 19–36.

———, ———, and J. Ohlson. 1992. Aggregate accounting earnings can explain most of security returns: The case of long event windows. *Journal of Accounting and Economics* 15: 119–142.

Elliott, J., and W. Shaw. 1988. Write-offs as accounting procedures to manage perceptions. *Journal of Accounting Research* 26 (Supplement): 91–119.

———, and J. Hanna. 1996. Repeated accounting write-offs and the information content of earnings. *Journal of Accounting Research* 34 (Supplement): 135–155.

Ely, K., and G. Waymire. 1999. Accounting standard-setting organizations and earnings relevance: Longitudinal evidence from NYSE common stocks, 1927–93. *Journal of Accounting Research* 37 (Autumn): 293–318.

Francis, J., and K. Schipper. 1999. Have financial statements lost their relevance? *Journal of Accounting Research* 37 (Autumn): 319–352.

Gigler, F., and T. Hemmer. 2001. Conservatism, optimal disclosure policy, and the timeliness of financial reports. *The Accounting Review* 76 (October): 471–493.

Givoly, D., and C. Hayn. 2000. The changing time-series properties of earnings, cash flows, and accruals: Has financial reporting become more conservative? *Journal of Accounting and Economics* 29: 287–320.

Gu, Z. 2002. Cross-sample incomparability of R^2s and additional evidence on value relevance changes over time. Working paper, Carnegie Mellon University.

Kim, M., and W. Kross. 2002. The ability of earnings to predict future cash flows has been increasing—Not decreasing. Working paper, University of Missouri–Columbia.

Kothari, S. P., and R. Sloan. 1992. Information in prices about future earnings: Implications for earnings response coefficients. *Journal of Accounting and Economics* 15: 143–171.

Lev, B., and P. Zarowin. 1999. The boundaries of financial reporting and how to extend them. *Journal of Accounting Research* 37 (Autumn): 353–385.

Liu, J., and J. Thomas. 2000. Stock returns and accounting earnings. *Journal of Accounting Research* 38 (Spring): 71–101.

Muller, K., and E. Riedl. 2001. Using stock returns to determine "bad" versus "good" news to examine the conservatism of accounting earnings. Working paper, Pennsylvania State University.

Pope, P., and M. Walker. 1999. International differences in timeliness, conservatism, and classification of earnings. *Journal of Accounting Research* 37 (Supplement): 53–99.

Ryan, S. 1988. Structural models of the accounting process and earnings. Ph.D. dissertation, Stanford University

———. 1995. A model of accrual measurement with implications for the evolution of the book-to-market ratio. *Journal of Accounting Research* 33 (Spring): 95–112.

———, and P. Zarowin. 1995. On the ability of the classical errors in variables approach to explain earnings response coefficients and R^2s in alternative valuation models. *Journal of Accounting, Auditing, and Finance* 10 (Fall): 767–786.

THE ACCOUNTING REVIEW
Vol. 78, No. 2
2003
pp. 555–580

The Effect of Measurement Alternatives on a Nonfinancial Quality Measure's Forward-Looking Properties

Karen L. Sedatole
The University of Texas at Austin

ABSTRACT: Drawing on theories from operations, management, and marketing, this paper examines whether three types of *measurement* alternatives affect the forward-looking properties of a quality measure. Using proprietary data from a medical services firm, the results show that the strength of the relation between a current quality measure and future quality-related warranty costs depends on: (1) the *standard* against which quality is measured, (2) assumptions regarding the relation's *symmetry*, and (3) assumptions regarding the relation's *functional form*. Thus, the *construction* of a quality performance measure may improve its ability to predict future financial performance and, hence, its usefulness for decision making and control. The paper concludes with an illustration of how firms can use a contemporaneous and forward-looking quality measure to assess quality-related decisions in a timely manner and shows that those decisions can be affected by the measurement alternatives examined.

Keywords: *leading indicator; nonfinancial performance measure; quality; cost of quality; cost function; service firms.*

Data Availability: *The author signed a confidentiality agreement that precludes revealing the identity of the firm that contributed data to this study or disseminating data without the firm's written consent.*

I. INTRODUCTION

Prior accounting research documents that quality-related nonfinancial performance measures such as customer satisfaction (Ittner and Larcker 1998a), defect rates, and on-time deliveries (Nagar and Rajan 2001) are often leading indicators of financial performance. Yet we know little about the factors that affect the strength of the relations between forward-looking quality measures and future financial performance (Ittner and

---

I am grateful for the assistance of my dissertation committee: Shannon Anderson (chair), David Butz, Raffi Indjejikian, Bill Lanen, and Leon Wyszewianski. I also thank the following individuals for their many helpful comments and suggestions: Ramji Balakrishnan, Shane Dikolli, Leslie Eldenburg, Steven Kachelmeier, Bill Kinney, Lisa Koonce, Joan Luft (associate editor), Mary Lea McAnally, Margaret Shackell-Dowell, Naomi Soderstrom, two anonymous reviewers, participants at the 2000 AAA/IMA Management Accounting Research Conference, the 2001 AAA Annual Meeting and workshop participants at The University of Texas at Austin, Stanford University, Rice University, University of North Texas, and The University of Texas at Dallas. I am also grateful to those individuals at my research site for their time and cooperation and for their efforts in providing data. This paper is based on my dissertation completed at the University of Michigan.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted June 2002
Accepted November 2002

Larcker 1998b; Lambert 1998). This paper examines whether *measurement* alternatives can strengthen the forward-looking properties of a quality measure and, therefore, its usefulness for decision making and control.

Forward-looking quality measures are important to managers because they link current quality-related decisions to future financial consequences. Conceptually, measure construction choices that strengthen the relation between a current quality measure and future financial performance increase the *congruence* of the measure to management objectives. The quality measure thus becomes potentially more useful for contracting (Datar et al. 2001; Feltham and Xie 1994). Moreover, current quality-related decisions are more likely to be value-maximizing when managers have a timely, yet highly congruent, quality measure by which to evaluate those decisions. The identification of measurement alternatives that can improve the forward-looking properties of a quality measure is, therefore, important to management accountants who play an important role within firms as sources of decision-useful information and as measurement experts (Kinney 2001).

Using a proprietary database of six units of a leading medical services firm, this study proposes and empirically examines whether three types of measurement alternatives affect the strength of the relation between a quality measure and future quality (i.e., warranty) costs. Drawing on theories from operations, management, and marketing to suggest measurement alternatives, the study examines: (1) the *standards* against which quality is measured, (2) the *symmetry* of the relation between the quality measure and future quality costs (i.e., whether variations in quality have a *uniform* affect on future quality costs), and (3) the *functional form* of the relation between the quality measure and future quality costs. The first speaks to the quality measure itself; the latter two speak to the form of the quality cost function.

There are three primary empirical results. First, the standard against which quality is measured affects the strength of the association between the quality measure and future quality costs. Marketing theory suggests that the customer's *ex ante* expectation of product performance or service outcome is the appropriate standard against which quality should be measured (e.g., Cronin and Taylor 1992, 1994; Parasuraman et al. 1985; Patterson 1993; Stank et al. 1999). However, in the current setting, an *absolute* quality measure (i.e., quality measured against an *optimal* outcome standard) is a better predictor of warranty costs than an *expectations-adjusted* quality measure (i.e., quality measured against an expectation standard).

Second, an asymmetry is documented in the quality (i.e., warranty) cost function in the current setting. This result is consistent with the generalizable hypothesis that the strength of the relation between the quality measure and future warranty costs is a function of customers' *expected net benefits* of warranty claims.

Finally, empirical results evaluating alternative functional forms of the warranty cost function do not support the tolerance limit approach that is commonly used in manufacturing settings and that assumes small variations from optimal quality are not costly for the firm. Instead, the results suggest that *all* deviations from optimal quality are costly in the current setting, consistent with quality theory in operations research (e.g., Taguchi et al. 1989). Moreover, the data support a linear (as opposed to an exponential) relation between the quality measure and future quality costs in the current setting.

The implication for management accounting of this study is that a quality measure can be made more useful for decision making and control by constructing the measure (i.e., with consideration of measurement alternatives such as those described above) in a way that improves its forward-looking properties. To highlight this, the paper concludes with an illustration of how firms can use an "improved" contemporaneous and forward-looking quality measure to assess quality-related decisions in a timely manner and shows that those decisions can differ when alternative forms of the quality measure are used.

Section II provides theoretical development for the quality measure and quality cost function measurement alternatives. Sections III and IV describe the research setting and the results of the empirical analysis, respectively. The decision-facilitating role of a quality measure is illustrated in Section V, and Section VI concludes.

II. THEORETICAL DEVELOPMENT

Prior research finds that quality measures (including customer satisfaction) are leading indicators of financial performance (e.g., Banker et al. 2000; Ittner and Larcker 1998a; Nagar 1998; Nagar and Rajan 2001). In this section, I draw from theories in operations, management, and marketing to examine the issues associated with measuring quality and with identifying the relation between a contemporaneous quality measure and future quality-related warranty costs (i.e., a quality cost function). First, I consider whether the choice of performance measure *standard* affects the strength of the relation between a quality measure and future quality costs. Second, I develop a hypothesis regarding the *symmetry* of the quality cost function (i.e., whether all variations in quality have a *uniform* affect on future quality costs). Finally, I consider alternative *functional forms* of the quality cost function.

Quality Standard

Quality standards are often absolutely determined. For example, precise product specifications are identified in manufacturing settings. Similarly, in a customer satisfaction response scale, the highest satisfaction level is the implicit absolute standard. Alternatively, an *expectation* of the quality outcome can be used as a standard against which a quality measure is constructed. For example, a firm may measure quality as the defect rate in a given period as compared to the historical average, possibly with exogenous, time-varying factors (e.g., volume) incorporated to further adjust for current period expectations.

Marketing theory suggests that a customer's overall satisfaction with a product or service (i.e., a measure of perceived quality) is a function of his *ex ante* expectation. That is, the gap between expectation and actual outcome determine the perceived quality (Cronin and Taylor 1992, 1994; Parasuraman et al. 1985; Patterson 1993; Stank et al. 1999). Marketing "gap analysis" documents a positive effect of perceived quality on firm profitability via repurchase (Oliver 1980; Cronin and Taylor 1992; Anderson and Sullivan 1993; Stank et al. 1999) and complaint behavior (Bearden and Teel 1983). In addition, there is some evidence that the relation between perceived quality of services and *overall* firm performance is positive (e.g., Phillips et al. 1983; Nelson et al. 1992; Anderson et al. 1994; Zeithaml et al. 1996).

Thus, marketing theory suggests that perceived quality measured relative to an expectation has forward-looking properties that may dominate an absolute quality measure. Since quality-related warranty work is precipitated by customers (note that this is ultimately true, even for product recalls), it follows that warranty costs are also likely to be a function of customer *ex ante* quality expectations. I, therefore, make the following hypothesis (stated in alternative form):

> **H1:** An *expectation-adjusted quality measure* (i.e., quality measured against the *ex ante* expected outcome) is more strongly associated with future quality costs (e.g., warranty costs) than an *absolute quality measure* (i.e., quality measured against the optimal outcome).

There are (at least) two potential reasons why an expectation-adjusted quality measure may *not* dominate an absolute quality measure. First, if customers fixate on optimal quality as a result of, for example, promotional promises (e.g., "if not completely satisfied" type statements) then, even though *ex ante* expectations may be for less than optimal quality, quality measured against the optimal outcome could be more highly associated with actual future warranty claims. Second, customers may have *ex ante* quality expectations that are too low; for example, a firm could purposefully set customer expectations low to increase *ex post* perceived quality. If this becomes evident to the customer *ex post*, then warranty costs may be more highly associated with an absolute quality measure once the customer updates his priors.

Asymmetry

Although prior research documents that quality-related performance measures are often associated with future financial performance (e.g., Banker et al. 2000; Ittner and Larcker 1998a; Nagar 1998; Nagar and Rajan 2001), theories predicting determinants of the strength of these relations have not been proposed in the accounting literature. This paper suggests one such theory, namely, that the strength of the relation between a quality measure and future quality-related warranty costs is increasing in the expected net benefits to the customer of the warranty work. That is, the expected net benefit moderates the quality-future cost relation. Although this hypothesized moderating effect is equally plausible for manufacturing firms, it is likely to be stronger in high-contact service firms for reasons described below.

An important characteristic of service firms that distinguishes them from traditional manufacturing firms is their "joint production." Many services (especially medical service firms) require a large degree of customer contact; that is, the service is jointly produced by both the customer and the firm's employee (Kotler and Bloom 1984; Parasuraman et al. 1985; Apte et al. 1997). Because of this characteristic, it is not *ex ante* clear that a contemporaneous quality measure will be a leading indicator of quality-related warranty services and costs. Joint production implies higher levels of customer participation in all service deliveries. As such, the *re-delivery* of an initial poor quality service imposes costs on service customers, including out-of-pocket costs and opportunity costs associated with the time it takes to participate in a second service delivery. Of course, the customer also has an *ex ante* expectation of the benefit of the warranty work. I hypothesize that the service customer's expected *net benefit* of warranty work is an important determinant of whether he or she pursues resolution of the poor initial service in the form of service re-delivery.

Expected net benefits may vary depending on characteristics of the customer. For example, a customer who expects to have future transactions with the firm may assess a higher expected net benefit to the resolution of a poor-quality service than a customer who does not. Thus, one might expect the relation between service quality and warranty costs to be stronger for repeat customers than for one-time customers.

More generally, situations in which a customer's cost of re-delivery of a poor service is relatively high (or benefits relatively low) result in a lower probability of the customer pursuing warranty work. This leads to the second measurement hypothesis to be tested, which is (stated in alternative form):

H2: The magnitude of the (negative) relation between quality and future warranty costs is *increasing in the expected net benefits* to the customer of the warranty work.

Stated another way, future warranty costs will be *more sensitive* to actual quality outcomes when the customer's expected net benefits of the warranty work are relatively higher.

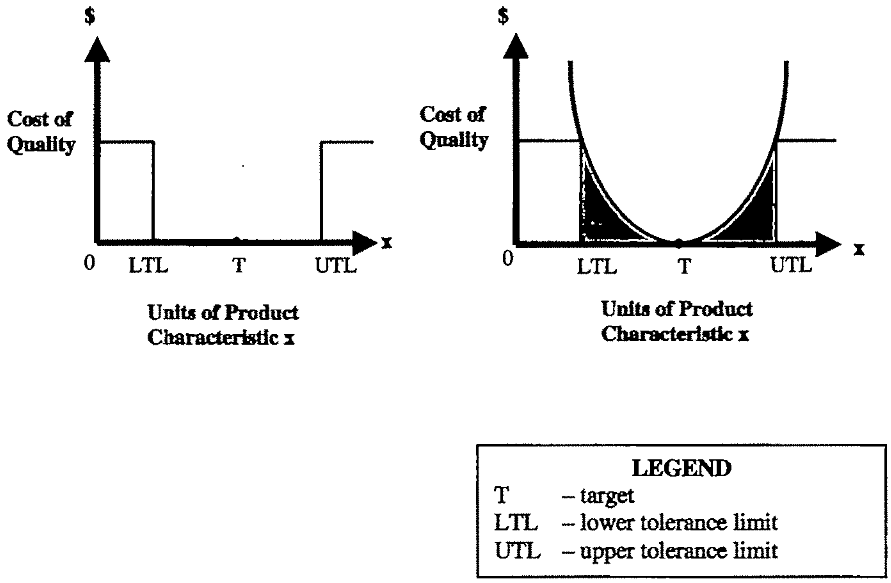
Functional Form

In any given setting, the relation between a measure of quality and future quality costs (i.e., the cost function) may take several functional forms, including (but not limited to) a dichotomous form as suggested by the traditional tolerance limit approach, a simple linear form, or even a quadratic form.

The traditional tolerance limit approach is widely used in product manufacturing practice and assumes that there is a range of quality outcomes that are equally acceptable. That is, as long as the product measurement outcome (i.e., a quality measure in manufacturing settings) is within predetermined "tolerance limits," regardless of how close it is to the target value, the product is deemed to be in conformance with quality standards (see Figure 1, Panel A). All products with dimensions outside of the upper and lower tolerance limits are nonconforming. The tolerance limit approach assumes that only nonconforming products result in nonconformance costs such as scrap, rework, and warranty costs (Albright and Roth 1992).

Quality management research in manufacturing settings indicates that the traditional tolerance limit approach of identifying good products is deficient because it ignores quality costs associated with small deviations from targets. Taguchi et al. (1989) argue that nonconformance quality costs will be incurred with *any* deviation of a product characteristic

FIGURE 1
Competing Quality Cost Function Theories


$
Cost of Quality
0
LTL
T
UTL
x
Units of Product Characteristic x
$
Cost of Quality
0
LTL
T
UTL
x
Units of Product Characteristic x
LEGEND
T – target
LTL – lower tolerance limit
UTL – upper tolerance limit

Panel A: The Traditional (Tolerance Limit) Model of Quality Costs

Panel B: Taguchi's Model of Quality Costs from its target (i.e., not just for products with characteristics outside of tolerance limits) (Figure 1, Panel B).[1] If this is true, then the cost function will assume a linear, or possibly quadratic, functional form.

Although Kim and Liao (1994) examine various functional forms of quality cost functions from a theoretical perspective, there is little work that attempts to discriminate empirically among these alternative forms. Identification of the correct functional form of a quality cost function provides a third measurement improvement that can potentially facilitate cost prediction and, ultimately, decision making and control. Since functional form is likely to be context-specific, I make no general predictions about the dominance of any particular functional form. Instead, in the empirical analysis I estimate and compare three functional forms: (1) a dichotomous cost function in which relatively small deviations from the quality standard are assumed to result in no increase in quality (i.e., warranty) costs, (2) a linear cost function in which *all* deviations are assumed to be linearly associated with increased costs, and (3) a quadratic cost function in which deviations are assumed to be associated with exponential increases in costs.

In summary, theories from the fields of operations, management, and marketing provide insight and suggest predictions regarding three quality measurement issues: (1) the alternative *standards* against which quality is measured, (2) whether variations in quality have a *uniform* affect on future quality costs (i.e., consideration of an asymmetric quality cost function), and (3) the exact *functional form* of the relation between the quality measure and future quality costs. The following two sections of the paper empirically examine the measurement issues described above.

III. RESEARCH SETTING AND DATA DESCRIPTION

Selection of the Research Site and Description of the Service

The research setting chosen for this study is a leading medical services firm that specializes in the surgical correction of myopic eyesight using excimer laser technology (hereafter, laser vision correction). This setting was chosen for several reasons. First, in contrast to the large amount of research on quality costs in manufacturing, there is little research of this type in service settings. The service setting chosen for this paper expands the quality research in a direction likely to be fruitful in the future (i.e., technology-driven service settings). Second, since the firm currently pursues a strategy of being a high-quality provider of laser vision correction, quality measurement issues studied are of particular importance.

Third, quality-related warranty (i.e., rework) costs associated with laser vision correction occur far in the future—over eight months in the future for the patients in the sample needing rework. Thus, a measure of quality contemporaneous to the surgical procedure would be useful to managers if shown to be a leading indicator of future costs. Finally, because of the technological nature of laser vision correction, the data are available to examine the measurement issues put forth in this paper. For example, the surgeons routinely identify an expected outcome distinct from the physiologically determined absolute quality standard (e.g., corneal curvature giving "20/20" eyesight). Indeed, a criticism of the marketing research previously discussed is the difficulty of measuring *ex ante* customer expectations, especially in service settings (Cronin and Taylor 1992; Parasuraman and Zeithaml 1994).

The sample consists of five surgery centers within one large firm. The centers are located in the U.S. and Canada and provide the surgical correction of varying levels of myopia (near-sightedness). The procedure involves surgically altering the shape of the patient's cornea using excimer lasers and the Photorefractive Keratectomy (PRK) procedure. The surgeries are performed at the center on an outpatient basis by an ophthalmologist who is a firm employee. The fee charged by the firm includes payment for an initial screening, the surgery, and all post-procedure exams performed by the patient's referring optometrist. The firm offers a lifetime guarantee and will provide, free of charge, an "enhancement" to patients so requiring. Thus, enhancements represent an important quality-related cost for the firm (i.e., warranty cost).

---

[1] Taguchi et al. (1989) argue that important sources of cost are the "hidden" losses of reputation damage and lost customer sales. They propose a quadratic form to account for these hidden losses.

Description of the Data

For each of the centers, there are two primary data sets (Table 1). The procedure-level outcomes database contains descriptive information for 4,470 initial (i.e., not enhancement) surgical outcomes. Table 1, Panel B presents the numbers of procedures performed and the number of months of data for each of the centers. Four of the five centers began operations during the period of study, May 1996 through September 1997.[2] Center A is the most established center, having the longest time-series of data (18 months) and the largest number of (initial) procedures (3,263) performed. The monthly level database contains financial information (investments in training and maintenance), center volume, and surgeon experience for 76 center-months.

Prior to the surgery, the surgeon takes two measurements of the patient's cornea—(1) a measure of the refractive spherical curvature of the cornea, the spherical measure; and (2) a measure of the shape of the cornea, the cylinder measure—each of which is measured in "diopters." The magnitude of the spherical measure (variable *MAG SPH*) indicates the degree of myopia, while the magnitude of the cylinder measure (*MAG CYL*) indicates the degree of astigmatism. Most procedures (80 percent in the sample) include both a spherical and a cylinder correction. The two measures are combined into one measure called the "spherical equivalent." This measure, *EQU*, is calculated as follows:

$$EQU = SPH + (CYL/2). \tag{1}$$

A spherical equivalent measure of 0.00 is called "*PLANO*." A *PLANO* measurement indicates no surgical correction is needed (i.e., light rays focus perfectly on the retina and there is no astigmatism) and is the optimal outcome, that is, an *absolute*, optimal quality standard. Note that the *EQU* measure is negative for myopic patients, with smaller (more negative) values indicating more severe corrections (see Figure 2).

The pre-surgery measurements, along with other factors such as patient age, are used to identify suitable candidates for the procedure and to determine the "intended" or expected spherical equivalent outcome (variable *E[outcome]*) used to test H1. Table 1, Panel A presents descriptive statistics for the variables used in the following two sections; Panel B presents the same statistics by center. Table 2 provides correlations (Spearman and Pearson) among the procedure level variables (Panel A) and the center-month variables (Panel B).[3]

---

[2] The original data consisted of ten centers. I omitted five centers because they entered the sample late in the test period and/or they had too few procedures during the test period. I allowed a center to enter the sample only when it reached a volume exceeding ten procedures in a month.

[3] During the period of study, there were no excimer lasers in the U.S. with FDA approval for the correction of hyperopia (i.e., spherical equivalent > 0). Moreover, no enhancements for *EQU* > 1.75 were performed in the Canadian centers. Since this externally imposed restriction would reduce the power of the statistical tests (or possibly bias the results), all observations for U.S. centers with *POST EQU* > 0 and Canadian centers with *POST EQU* > 1.75 were eliminated from the analysis.

TABLE 1
Descriptive Statistics

Panel A: Descriptive Statistics[a]

No. of Center-Months	76
No. of Initial Procedures	4,470
No. of Enhancements[b]	243
Enhancement Rate[c]	5.44%

Procedure-Level Variables	**Mean**	**Std. Dev.**	**Min.**	**Max.**
DURATION (with censored observations)	265.52	148.63	11	546
DURATION (without censored observations)	255.48	71.91	118	491
DEVIATION (quality measure)	0.26	0.45	0	4.38
MAG SPH (magnitude of the *SPH* correction)	4.63	2.62	0.00	18.00
MAG CYL (magnitude of the *CYL* correction)	1.06	0.91	0.00	6.00
AGE	38.20	8.98	17	68
GENDER (1 = male)	0.50	0.50	0	1
PRE EQU	−5.16	2.69	−19.25	−0.63
POST EQU	−0.08	0.66	−4.38	4.63
Monthly-level variables				
EXPERIENCE[d]	943.53	1,448.00	11	5,102
TRAINING (1 = post-July 97)	0.13	0.34	0	1
Avg *MAINT*	15.17	20.53	2.31	51.62
VOLUME	81.97	88.91	6	309

Panel B: Descriptive Statistics by Center[e]

Center:	**A**	**B**	**C**	**D**	**E**	**All Centers**
No. of Months	18	15	14	13	16	76
No. of Initial Procedures	3,263	290	393	186	338	4,470
No. of Enhancements[b]	202	3	25	4	9	243
Enhancement Rate[c]	6.19%	1.03%	6.36%	2.15%	2.66%	5.44%
Procedure-Level Variables						
DURATION	299.02	190.15	167.01	119.56	201.60	265.52
(with censored observations)	(142.47)	(136.81)	(109.02)	(105.96)	(132.26)	(148.63)
DURATION	265.00	301.00	196.20	213.50	209.89	255.48
(without censored observations)	(68.96)	(122.65)	(68.90)	(36.82)	(34.34)	(71.91)
DEVIATION	0.21	0.41	0.43	0.28	0.43	0.26
(quality measure)	(0.39)	(0.52)	(0.66)	(0.43)	(0.51)	(0.45)
MAG SPH	4.88	3.42	4.32	3.51	4.23	4.63
(magnitude of the *SPH* correction)	(2.74)	(1.78)	(2.32)	(2.04)	(2.08)	(2.62)
MAG CYL	1.15	0.66	0.93	0.70	0.91	1.06
(magnitude of the CYL correction)	(0.92)	(0.68)	(0.83)	(0.85)	(0.87)	(0.91)
AGE	38.28	36.47	38.21	38.95	38.42	38.20
	(8.98)	(8.51)	(9.02)	(8.65)	(9.43)	(8.98)
GENDER (1 = male)	0.50	0.47	0.49	0.57	0.49	0.50
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
PRE EQU	−5.46	−3.75	−4.77	−3.86	−4.69	−5.16
	(2.80)	(1.81)	(2.36)	(1.98)	(2.09)	(2.69)
POST EQU	0.04	−0.41	−0.43	−0.28	−0.43	−0.08
	(0.66)	(0.52)	(0.66)	(0.43)	(0.51)	(0.66)

(continued on next page)

TABLE 1 (continued)

Center:	A	B	C	D	E	All Centers
Monthly-level variables						
EXPERIENCE[d]	3,275.28	221.87	262.50	120.15	261.75	943.53
	(1,267.47)	(157.21)	(197.89)	(95.14)	(191.95)	(1,448.00)
TRAINING (1 = post-July 97)	0.11	0.13	0.14	0.15	0.13	0.13
	(0.32)	(0.35)	(0.36)	(0.38)	(0.34)	(0.34)
Avg *MAINT*	51.62	2.91	2.31	7.88	2.81	15.17
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(20.53)
VOLUME	230.50	34.60	45.57	24.23	38.06	81.97
	(48.21)	(26.29)	(26.56)	(16.55)	(22.79)	(88.91)

---

[a] For the quality measure, *DEVIATION*, the quality cost proxy measure, *DURATION*, and the procedure-level variables (*MAG SPH*, *MAG CYL*, *AGE*, *GENDER*, *PRE EQU*, and *POST EQU*), the statistics are computed over all outcomes. For the monthly-level variables (*VOLUME*, *TRAINING*, Avg *MAINT*, and *EXPERIENCE*), the statistics are computed over all center-months.

[b] The number of initial procedures in the database that required an enhancement during the test period.

[c] Enhancement rate is the percentage of initial procedures that required at least one enhancement. Sixteen patients required two enhancements. Note that this is a right-censored statistic.

[d] The experience variable is computed using an additional 11 months of outcomes data for which there was no corresponding monthly data.

[e] Standard deviations in parentheses.

Variable Definitions:

DURATION = the number of days between the initial procedure and the date of the enhancement (or to the end of the test period for censored observations). This variable proxies for future quality costs;

DEVIATION = the difference between the actual spherical equivalent outcome and an outcome of *PLANO* (i.e., spherical equivalent equal to zero) for observations in which *POST EQU* < 0 and zero otherwise. This variable is the contemporaneous nonfinancial quality measure;

MAG SPH = the magnitude of the spherical equivalent correction needed (i.e., the severity of the myopia);

MAG CYL = the magnitude of the cylindrical correction needed (i.e., the severity of the astigmatism);

AGE = patient age;

GENDER = gender of patient (1 = male);

PRE EQU = the patient's spherical equivalent measure (in diopters) just prior to the procedure;

POST EQU = the patient's spherical equivalent measure (in diopters) after the procedure;

EXPERIENCE = the cumulative number of procedures;

TRAINING = indicator variable identifying the center-months following the firm-wide training program (1 = post-July 97);

Avg *MAINT* = maintenance expenditure per procedure for the center; and

VOLUME = the number of procedures performed in a center-month.

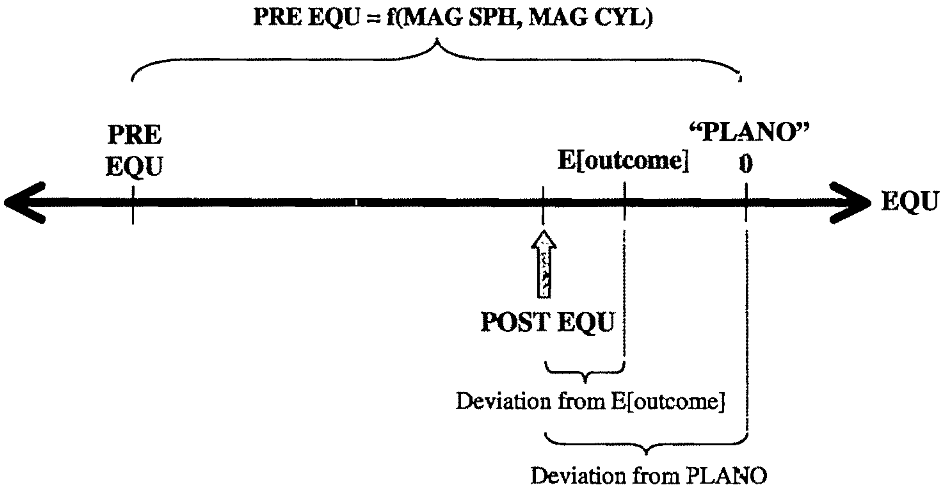
IV. QUALITY AND QUALITY COST FUNCTION MEASUREMENT

In this section, following a brief discussion of the warranty cost proxy used in the tests, I empirically examine the three types of quality measurement alternatives identified in Section II.

Warranty Cost Proxy

A significant source of quality-related costs in the laser vision correction industry is the cost of performing "enhancements" (i.e., rework) for those patients who require them. I argue that enhancement rates capture three dimensions of quality-related costs. First, there are out-of-pocket (variable) costs of performing these procedures including the cost of medications, technical support staff, and laser fees. Second, reputation effects resulting from dissatisfied customers who required an enhancement may impose costs on the firm in the

FIGURE 2
Relations among Key Variables



LEGEND

PRE EQU	– spherical equivalent measure prior to the procedure
MAG SPH	– the magnitude of the required spherical correction
MAG CYL	– the magnitude of the required cylindrical correction
POST EQU	– actual spherical equivalent measure after the procedure
PLANO	– optimal spherical equivalent outcome; represents the absolute (optimal) quality standard
E[outcome]	– the expected spherical equivalent outcome identified by surgeon just prior to the procedure; represents the expectation quality standard

form of lost referrals. Finally, enhancements impose opportunity costs because of the inability to perform paid procedures while the enhancement procedure is being performed—a cost that is significant for centers operating at or near capacity (e.g., Center A).

The current analysis is limited by the lack of data on the explicit costs of enhancements. However, out-of-pocket costs vary little across enhancement procedures. It is, therefore, unnecessary to attempt to impute an out-of-pocket dollar cost to each enhancement. Rather, warranty cost is effectively captured by a dichotomous variable of the incidence of an enhancement. Moreover, interviews conducted with surgeons and staff optometrists reveal that more severe quality "failures" result in *earlier* enhancements, and that the probability

TABLE 2
Correlations

Panel A: Procedure-Level Variables[a]

Spearman Rank Correlations

	DURATION	*DEVIATION*	*MAG SPH*	*MAG CYL*	*AGE*	*GENDER*	*PRE EQU*	*POST EQU*
DURATION		0.004	0.087*	0.082*	0.003	0.018	−0.100*	0.071*
DEVIATION	0.011		−0.007	−0.065*	0.056*	−0.057*	0.015	−0.924*
MAG SPH	0.081*	0.018		0.068*	0.049*	−0.096*	−0.981*	0.062*
MAG CYL	0.079*	−0.040*	0.057*		0.132*	−0.014	−0.234*	0.096*
AGE	0.005	0.087*	0.044*	0.095*		−0.021	−0.069*	−0.028
GENDER	0.017	−0.062*	−0.083*	−0.008	−0.021		0.099*	0.048*
PRE EQU	−0.093*	−0.011	−0.986*	−0.224*	−0.059*	0.083*		−0.075*
POST EQU	0.058*	−0.838*	0.100*	0.080*	−0.038	0.045*	−0.111*	

Pearson Correlations

Panel B: Monthly Level Variables[a]

Spearman Rank Correlations

	EXPERIENCE	ln *EXPERIENCE*	*TRAINING*	Avg *MAINT*
EXPERIENCE		1.000*	0.319*	0.431*
ln *EXPERIENCE*	0.858*		0.319*	0.431*
TRAINING	0.125	0.231		−0.018
Avg *MAINT*	0.896*	0.810*	−0.032	

Pearson Correlations

---

* Indicates significance at the 0.01 level.

[a] Both Pearson (lower diagonal) and Spearman (upper diagonal) correlations are provided.

The variables are defined in Table 1.

that an enhancement will be given is decreasing in the amount of time that has elapsed (i.e., without an enhancement).[4] For these reasons, the *incidence* and *timing* of enhancements is a reasonable proxy for quality-related warranty costs in the current setting.

Based on the arguments given above, I use the duration of the initial procedure, calculated as the number of days from the initial procedure to the enhancement (or the end of the test period if no enhancement is performed), as the warranty cost proxy.[5] Note that *DURATION* is inversely related to costs; that is, high values of *DURATION* indicate a low likelihood of enhancement. For those patients requiring enhancements, the minimum, mean, and maximum *DURATION* are 118, 255, and 491, days, respectively (see Table 1, Panel A).

Comparison of Absolute and Expectation-Adjusted Quality Standards

Marketing theory suggests that the *ex ante* expected outcome is a more appropriate standard for evaluating quality performance than is an absolute, optimal standard. Hypothesis 1 accordingly predicts that an *expectation-adjusted* quality measure is more strongly associated with future quality costs (i.e., warranty costs) than an absolute quality measure. In the current setting, *PLANO* and *E[outcome]* represent the absolute (i.e., optimal) and expectation standards, respectively. Thus, the difference between the initial post-surgery spherical equivalent measure, *POST EQU*, and *PLANO* is an *absolute* quality measure. Similarly, the difference between *POST EQU* and *E[outcome]* is an *expectation-adjusted* quality measure.

I test H1 by comparing the strength of the associations between each of these measures and the warranty cost proxy, *DURATION*. A maximum likelihood estimation of procedure duration, *DURATION*, as a function of each measure is estimated. This estimation assumes a Weibull distribution for duration and adjusts for right censoring of procedures not requiring an enhancement by the end of the test period.[6]

The results of the test of H1 are presented in Table 3. Both the absolute and the expectations-adjusted measures (Models 1 and 2, respectively) are significantly (p-value $<$ .001) negatively associated with *DURATION*, indicating that larger deviations from the respective standard (i.e., lower quality) result in a shorter procedure duration—that is, a higher likelihood of a future enhancement (i.e., higher future warranty costs). However, the model with the absolute quality measure has a pseudo-R^2 of 6.63 percent, while the model with the expectation-adjusted measure has a pseudo-R^2 of only 2.31 percent. Moreover, when both measures are included in the model, only the absolute measure remains significant; that is, the absolute measure is incrementally informative over the expectation-adjusted measure, but not vice versa. This suggests that, *contrary to H1* and prior theory,

---

[4] Although I do not have data to corroborate this claim, it also seems reasonable to assume that hidden quality costs (Taguchi et al. 1989) related to customer dissatisfaction and reputation damage are also decreasing in elapsed time.

[5] Modeling the duration of the initial procedure also offers a significant econometric advantage. Because enhancements occur, on average, over eight months following the initial procedure; and because my sample includes a fairly short time series, the data are significantly right-censored. A duration modeling approach adjusts for right censoring, allowing the information in *both* censored and uncensored observations to be used in a way that provides consistent estimates of the parameters. In fact, this is one of the most significant advantages of the survival analysis technique (Allison 1995).

[6] The Weibull distribution allows for either a monotonically increasing or decreasing hazard function (Greene 1997) and is less restrictive than other nonnegative distributions often used in duration models (e.g., exponential or lognormal distributions). With this specification, each coefficient estimate represents the derivative of the log of *E[duration]* with respect to a unit change in the associated independent variable (i.e., $\partial \ln E[DURATION] / \partial x_k = \beta_k$). The estimated Weibull distribution scale parameter, γ, indicates the duration dependence of the data.

TABLE 3
Comparison of the Optimal and Expectation-Adjusted Quality Standards[a]

Dependent Variable:		**DURATION**	
Independent Variable	**Model 1**	**Model 2**	**Model 3**
Intercept	7.815***	7.701***	7.829***
Deviation from *PLANO*	−0.500***		−0.428***
Deviation from *E[outcome]*		−0.497***	−0.151
γ (scale parameter)	0.638	0.669	0.632
Pseudo-R^2	6.63%	2.31%	6.15%
LR Statistics:			
vs. intercept only model	18.42***	7.72***	19.22***
No. of noncensored values	36		
No. of right censored values	277		

*** Indicates significance at the .001 level (two-tailed test).

[a] This table presents the results of a maximum likelihood estimation of procedure *DURATION* (i.e., time to enhancement) as a function of quality as measured against one of two standards, the optimal outcome of *PLANO* (i.e., the absolute quality measure) or the expected outcome (i.e., the expectation-adjusted quality measure). The model is estimated for the subsample of observations in which *POST EQU* < 0 and *E[outcome]* ≠ 0 (*PLANO*). The estimation assumes a Weibull distribution for duration and adjusts for right censoring for procedures not requiring an enhancement by the end of the test period.

Coefficient estimates are shown along with estimates of the Weibull distribution scale parameter, γ.

Two fit statistics, the pseudo-R^2 and the likelihood ratio statistic, are also presented.

Variable Definitions:

DURATION = the number of days between the initial procedure and the date of the enhancement (or to the end of the test period for censored observations);

PLANO = the optimal spherical equivalent outcome; represents the optimal quality standard and assumes the value of 0;

E[outcome] = the expected spherical equivalent outcome identified by surgeon just prior to the procedure; represents the expectation quality standard; and

POST EQU = the actual spherical equivalent measure after the procedure.

Deviation from *PLANO* (i.e., Absolute Quality Measure) = *Abs(POST EQU − PLANO)*

Deviation from *E[outcome]* (i.e., Expectation-Adjusted Quality Measure) = *Abs(POST EQU − E[outcome])*

quality measured against the ***absolute*** standard is more highly associated with future warranty cost in the current setting than is quality measured against the *expectation* standard.

One potential explanation for this somewhat surprising finding is that customers, regardless of expectations the surgeons attempt to set, fixate on the optimal outcome of 20/20 vision and implicitly use *PLANO* as the standard. In other words, the surgeon's expected outcome, although communicated to the patient, was not the patient's expected outcome. Consistent with the previously described criticism of extant marketing gap analysis research, this highlights the difficulty and importance of measuring *ex ante* expectation (an additional measurement issue, but one that is beyond the scope of this paper), a common concern in the marketing gap analysis research described above.

A second potential explanation for the unexpected result is the surgeon's incentive to set an overly conservative expectation standard to increase the likelihood of exceeding this standard. Although medical professionals are, by law, restricted from being compensated based on medical outcomes, the surgeons likely have implicit incentives (e.g., reputation) to provide high quality. Because the expected outcome is determined subjectively by the surgeon, he has both the incentive *and* the opportunity to understate the expected quality. The above-described findings are consistent with an overly conservative estimate of the expected quality outcome on the part of the surgeon, and with the warranty claim (i.e., enhancement) decision ultimately being made based on the optimal standard of *PLANO*.

Asymmetry in the Quality Cost Function

The results presented above document that procedure *DURATION* is decreasing in (an *absolute* measure of) quality. In this section, I examine whether the magnitude of this relation is uniform, in this case, across customers. Hypothesis 2 predicts that the magnitude of the negative relation between quality and warranty costs is increasing in the expected net benefits to the customer of the warranty work. The current setting provides a natural means of testing this hypothesis.

Recall that the original procedures in this sample were for the correction of myopia in which the cornea is too curved and must be flattened in the center by the laser. Surgical outcomes to the right of *PLANO* (see Figure 2) represent *over*corrections by the surgeon; the patient is now hyperopic with a cornea that is too flat. Since tissue cannot be added to the cornea to increase the curvature back to *PLANO*, an "enhancement" would require the surgeon to laser around the perimeter of the cornea to induce curvature. This is a more difficult correction with a lower likelihood of success, in part because each individual has a limited amount of corneal tissue that can be removed. I assert that an initial overcorrection by the surgeon results in lower expected net benefits to the customer of an enhancement (i.e., warranty work). Thus, H1 predicts that the negative relation between *DURATION* and quality, as measured by deviations from the *PLANO* standard, will be weaker for initial procedures that result in an overcorrection (i.e., $POST\ EQU > PLANO$) relative to the outcome of an undercorrection (i.e., $POST\ EQU < PLANO$).

I test this hypothesis using a maximum likelihood estimation of procedure duration, *DURATION*, as a function of deviations from the *PLANO* standard and a term interacting this variable with the overcorrection indicator variable. This estimation again assumes a Weibull distribution for *DURATION* and adjusts for right censoring in the data.

The results, presented in Table 4, indicate that the magnitude of deviations from the target outcome of *PLANO* is negatively associated with *DURATION* for undercorrections (coefficient of -0.439, p-value $< .001$). That is, for actual outcomes less than *PLANO*, the further the actual outcome is from the intended outcome of *PLANO*, the shorter the duration of the initial procedure (i.e., warranty work is more likely). However, the negative relation between deviations from *PLANO* and *DURATION* disappears for overcorrections as evidenced by the significantly positive interaction term (p-value $< .001$). In fact, the sum of the two coefficient estimates is significantly positive (p-value $< .001$), indicating that larger overcorrections result in longer procedure durations (i.e., increasingly *unlikely* that a enhancement will be given). Thus, the evidence is consistent with H2, which predicts that a decrease in the expected net benefits to the customer of warranty work weakens the negative relation between quality and future warranty costs.

The results described above should not be interpreted to mean that overcorrections are costless and should, therefore, become the objective. Overcorrections, although not associated with increased *warranty costs*, may well be very costly for the firm in terms of customer dissatisfaction and reputation damage. A limitation of this study is the lack of data to proxy for these costs (e.g., customer satisfaction data) and the necessary restriction of the analysis to actual warranty costs. Still, the analysis illustrates that the sensitivity of future warranty costs to a quality measure may by moderated by customers' cost-benefit assessments of the warranty work. The extent to which this generalizes to other settings is a question for future research.

TABLE 4
Test for Asymmetry in the Quality Cost Function[a]

Independent Variable	Dependent Variable: DURATION
Intercept	7.574***
Deviation from *PLANO*	−0.439***
Deviation from *PLANO* * *OVER*	0.835***
γ (scale parameter)	0.555
Pseudo-R^2	5.45%
LR Statistics:	
vs. intercept only model	122.24***
vs. model without interaction	63.62***
No. of noncensored values	280
No. of right censored values	4,538
No. of undercorrections (mean absolute deviation)	2,400 (0.563)
No. of overcorrections (mean absolute deviation)	2,418 (0.572)

*** Indicates significance at the .001 level (two-tailed tests).

[a] This table presents the results of a maximum likelihood estimation of procedure *DURATION* (i.e., time to enhancement) as a function of quality as measured against the optimal (*PLANO*) standard and an interaction of this variable and an overcorrection indicator variable, *OVER*. The estimation assumes a Weibull distribution for duration and adjusts for right censoring for procedures not requiring an enhancement by the end of the test period.

Coefficient estimates are shown along with estimates of the Weibull distribution scale parameter, γ.

Two fit statistics, the pseudo-R^2 and the likelihood ratio statistic, are also presented.

Variable Definitions:

DURATION = the number of days between the initial procedure and the date of the enhancement (or to the end of the test period for censored observations);

PLANO = the optimal spherical equivalent outcome; represents the optimal quality standard and assumes the value of 0;

POST EQU = the actual spherical equivalent measure after the procedure; and

OVER = an indicator variable equal to one for overcorrections.

Deviation from *PLANO* standard (i.e., Absolute Quality Measure) = *Abs*(*POST EQU* − *PLANO*)

Functional Forms of the Quality Cost Function

The final measurement test is of the functional form of the quality (i.e., warranty) cost function. The empirical tests in this section involve estimating and comparing the fits of three models, distinguished only by the three functional form assumptions: the dichotomous form representing the traditional tolerance limit approach, and two forms (linear and quadratic) representing variations of the Taguchi et al. (1989) cost function theory.

Recall that the preceding analysis reveals an asymmetry in the warranty cost function, namely, that warranty costs are *not* increasing in the magnitude of *overcorrections*. I, therefore, restrict my functional form tests to the subsample of undercorrected procedures.

To distinguish between the three proposed alternative functional forms of the cost function, I construct three measures of quality, all of which are based on deviations of actual surgery outcomes, *POST EQU*, from the *absolute* standard of *PLANO* (*EQU* = 0). The first measure, *OUTTOL*, is constructed to model the traditional tolerance limit approach that assumes quality costs are zero for outcomes within a tolerance limit around the standard. The tolerance limit is arbitrarily chosen as 1 diopter. The second and third measures reflect the theory that *any* deviation from the outcome standard is costly (Taguchi et al. 1989).

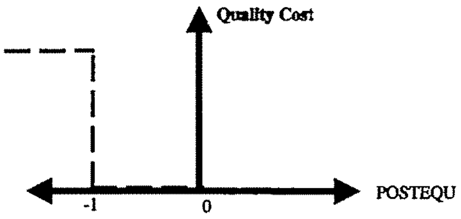
These measures, *DEVIATION* and *DEVIATION*^{2}, assume a linear and quadratic form, respectively. The three measures are:

(1) *OUTTOL*, representing the traditional tolerance limit cost function (Figure 3, Panel A) and defined as a dichotomous variable equal to 1 if *Abs*(*POST EQU*) > 1.0 diopter, and 0 otherwise;
(2) *DEVIATION*, representing the linear cost function (Figure 3, Panel B) defined as *Abs*(*POST EQU*); and
(3) *DEVIATION*^{2}, representing the quadratic cost function (Figure 3, Panel C) defined as [*Abs*(*POST EQU*)]^2

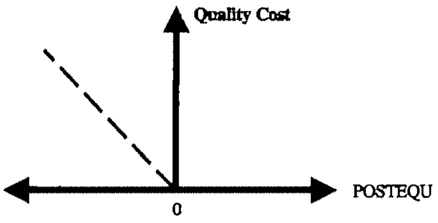
where *POST EQU* is the spherical equivalent measure following the procedure.

FIGURE 3
Alternative Functional Forms of the Quality Cost Function

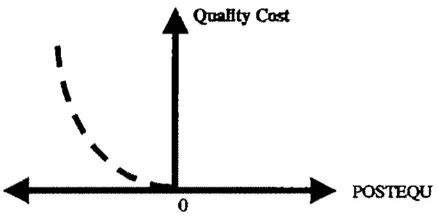
Panel A: Dichotomous (Tolerance Limit) Quality Cost Function



Panel B: Linear Quality Cost Function



Panel C: Quadratic Quality Cost Function



Three models are estimated in which quality cost (proxied by *DURATION*) is modeled separately as a function of each of these variables.

Table 5 presents the results of the maximum likelihood estimation of quality costs, proxied by procedure duration, as a function of each of the three measures (Model 4 includes all three measures). I again assume a Weibull distribution for *DURATION* and adjust for right censoring.

Consistent with expectations, all three measures are significantly negatively associated with *DURATION*, indicating that larger deviations from a *PLANO* outcome result in shorter initial procedure duration (i.e., enhancement more likely). However, in a model with all three measures (Model 4), *DEVIATION* retains its significance (and sign) while the other two measures do not (p-value > .10). Thus, I am able to discriminate between the three non-nested models and accept the *DEVIATION* (linear) model over the other two (Darnell 1994).

In summary, I find that, using the optimal outcome of *PLANO* as the quality standard, neither the dichotomous measure (i.e., representing the tolerance limit approach traditionally used in manufacturing), nor the measure based on squared deviations of actual outcomes from the quality standard (i.e., representing the quadratic functional form) add explanatory power over the linear measure's ability to predict quality costs.

Summary of Measurement Improvements

Evidence provided in the tests described above suggests that the standard against which a quality measure is constructed may affect the measure's ability to predict future financial performance. The test of H1 documents that the absolute, optimal outcome of *PLANO* is the appropriate standard for measuring quality in the current setting. That is, a measure constructed relative to this standard has a stronger association with future warranty costs than a quality measure using the expected outcome as the standard. Hypothesis 2 suggests an asymmetry in the relation between the nonfinancial performance measure and future warranty costs; namely, the strength of the quality-cost relation is increasing in the customer's expected net benefits of warranty work. Finally, the previous section confirms that all deviations from the standard are costly for the firm, consistent with Taguchi et al. (1989). Moreover, a linear quality measure dominates the quadratic measure suggesting a linear quality (warranty) cost function. These findings, taken together, suggest the following "improved" quality measure:

$$DEVIATION = \begin{cases} Abs(POST\ EQU) & \text{if } POST\ EQU < 0 \\ 0 & \text{otherwise} \end{cases}$$

where *POST EQU* is the spherical equivalent measure following the procedure. Descriptive statistics for this variable are reported in Table 1. Admittedly, the *specific* findings presented in the preceding three subsections are not generalizable beyond the current setting. However, three measurement issues have been identified—(1) performance measure standard, (2) assumed symmetry of the cost function, and (3) assumed functional form of the cost function—and shown to affect the strength of the relation between the nonfinancial quality measure and future costs. These measurement issues and the theories presented to suggest the measurement alternatives are likely to generalize beyond the measurement of quality and beyond the current setting.

TABLE 5
Test of Alternative Functional Forms of the Quality Cost Function[a]

Dependent Variable:	**DURATION**			
	Model 1 Dichotomous Model	**Model 2 Linear Model**	**Model 3 Quadratic Model**	**Model 4**
Independent Variable				
Intercept	7.305***	7.465***	7.295***	7.505***
OUTTOL	−0.371***			0.226
DEVIATION		−0.355***		−0.487***
DEVIATION^2			−0.093***	0.008
γ (scale parameter)	0.569	0.565	0.562	0.565
Pseudo-R^2	0.82%	2.13%	1.65%	1.99%
LR Statistics:				
vs. intercept only model	13.92***	32.95***	25.92***	34.86***
vs. Model 2				1.91
No. of noncensored values	211			
No. of right censored values	2,165			

*** Indicates significance at the .001 level (two-tailed tests).

[a] This table presents the results of a maximum likelihood estimation of procedure *DURATION* (i.e., time to enhancement) as a function of one of the three quality measures: *OUTTOL*, *DEVIATION*, or *DEVIATION*^{2} (see Figure 3). The model is estimated for the subsample of observations in which *POST EQU* < 0. The estimation assumes a Weibull distribution for duration and adjusts for right censoring for procedures not requiring an enhancement by the end of the test period.

Coefficient estimates are shown along with estimates of the Weibull distribution scale parameter, γ.

Two fit statistics, the pseudo-R^2 and the likelihood ratio statistic, are also presented.

Variable Definitions:

DURATION = the number of days between the initial procedure and the date of the enhancement (or to the end of the test period for censored observations);

OUTTOL = represents the traditional tolerance limit cost function (Figure 3, Panel A); defined as a dichotomous variable equal to 1 if *Abs*(*POST EQU*) > 1.0 diopter, and 0 otherwise;

DEVIATION = represents the linear cost function (Figure 3, Panel B); defined as *Abs*(*POST EQU*); and

DEVIATION^2 = represents the quadratic cost function (Figure 3, Panel C) defined as $[Abs(POST\ EQU)]^2$.

V. QUALITY MANAGEMENT ILLUSTRATION

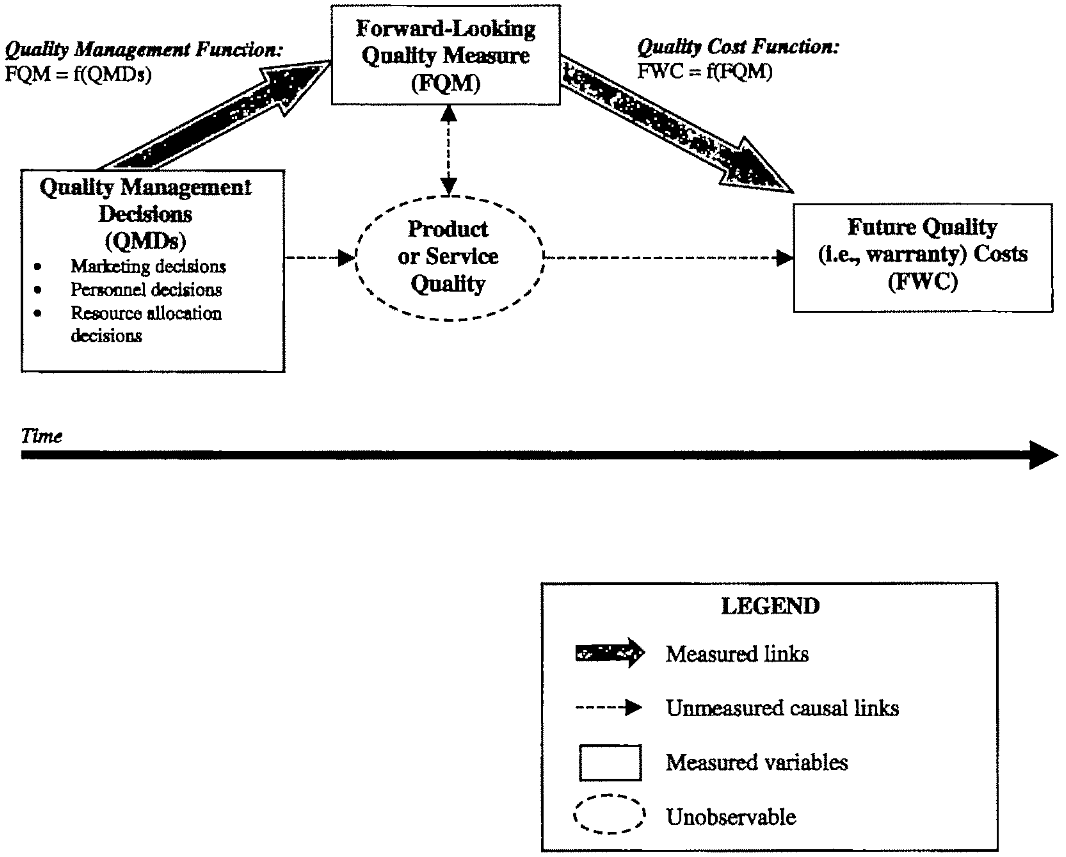
Implications for Decision Making

Forward-looking quality measures such as the one constructed above play an important role in understanding the relation between quality management decisions and future quality costs (see Figure 4). Management makes numerous quality-related management decisions. Often, there is a significant time lag between these decisions and the incurrence of quality costs such as warranty costs. As a result, a quality measure constructed contemporaneous to the service delivery and with forward-looking properties can play an important role as a connecting link between management decisions and future quality costs.

Improving the measurement of quality and of the quality cost function (such as was done in the preceding section) is the first important role that management accounting can play in quality management. The second is the linking of quality back to the management decisions that drive quality (and, ultimately, quality costs). A "quality management function" that empirically ties management decisions to a quality measure with forward-looking properties provides *timely* management feedback for improved decision making and control.

As an illustration, a quality management function (with the "improved" quality measure, *DEVIATION*, as the dependent variable) is estimated for the current setting. Seven

FIGURE 4
Quality Cost and Quality Management Functions


Quality Management Function:
FQM = f(QMDs)
Forward-Looking Quality Measure (FQM)
Quality Cost Function:
FWC = f(FQM)
Quality Management Decisions (QMDs)
Marketing decisions
Personnel decisions
Resource allocation decisions
Product or Service Quality
Future Quality (i.e., warranty) Costs (FWC)
Time
LEGEND
Measured links
Unmeasured causal links
Measured variables
Unobservable

quality-related decisions (i.e., quality drivers) are identified and represent three categories of quality management decisions likely to be important in service settings (see Table 6 for examples). The first category, *marketing decisions,* include the gender and age of the patient and the magnitudes of required spherical and cylindrical corrections. These variables, and especially the latter two, represent service complexity. Prior accounting research finds product and process complexity to be significant drivers of costs—including quality costs—in manufacturing (Anderson 1993, 1995; Banker and Johnston 1993; Datar et al. 1993; Ittner and MacDuffie 1995; MacArthur and Stranahan 1998) and in healthcare settings (Balakrishnan et al. 1996). Managers of service firms must choose the complexity level of services to provide, leading to the selection of these variables as important quality-related marketing decisions.[7]

---

[7] In some settings, the firm has a large degree of control over which customers will receive services (e.g., healthcare settings in which the physician can "screen" patients for a particular treatment). In other settings, although the degree of control is not as great, the firm influences customer selection by its pricing and advertising decisions.

TABLE 6
Quality-Related Management Decisions in Service Firms

Type of Decision	Decision Examined in Current Study
Marketing Decisions	
• complexity of service delivered	→ magnitudes of spherical and cylindrical corrections
• heterogeneity of services delivered (i.e., service mix)	
• location of service delivery	
• other customer-specific characteristics (individual or firm)	→ gender and age of the patient
Personnel Decisions	
• hiring—prior experience	→ surgeon experience
• hiring—level and quality of prior training/education	
• hiring—level of risk aversion	
• incentive system design	
• performance measurement system design	
Resource Allocation Decisions	
• investments in training	→ investments in surgeon and technician training
• investments in maintenance	→ investments in laser equipment maintenance
• investments in appraisal activities	
• firm-wide quality initiatives (e.g., TQM)	

Personnel decisions also affect quality and include the experience of the employee delivering the service (e.g., surgeon experience). Learning occurs with experience (or over time) and leads to improved quality (e.g., Fine 1986; Ittner et al. 2001). The importance of personnel decisions will likely be increasing in the extent of interaction between the customer and the employee.

Finally, firms must make *resource allocation decisions* such as investments in training and equipment maintenance. Prior research documents significant relations between investments in conformance activities and quality (nonconformance) costs in manufacturing firms (e.g., Datar et al. 1993; Ittner 1996; Ittner et al. 2001).[8]

Quality Management Function

Table 7, Model 1 presents the results of an OLS estimation of *DEVIATION* as a function of the seven quality-related management choice variables. Two control variables are included, an indicator variable for Center A (to control for differences in this center not captured by the other variables) and an indicator variable for overcorrections.[9]

---

[8] The traditional view of cost of quality management assumes that nonconformance costs can be reduced only by increasing conformance costs (Feigenbaum 1983; Ireson 1971). More recent theories suggest that, in a dynamic setting, both conformance and nonconformance costs can be reduced with organizational learning (Crosby 1979; Fine 1986; Love et al. 1995) and with capital investments in quality initiatives that allow for continual quality improvement. Although empirical evidence supports the dynamic model (e.g., Foster and Adam 1996; Ittner 1996; Ittner et al. 2001), the current study does not capture this dynamic environment.

[9] Equivalent results are obtained for a subsample of undercorrections.

TABLE 7
Estimation of the Quality Management Function[a]

Dependent Variable:		**Model 1** ***DEVIATION***	**Model 2** ***DEVIATION***^{2}	**Model 3** ***OUTTOL***
Independent Variable	**Exp. Sign**			
Intercept		−0.298*	−1.150***	−7.458***
Marketing Decisions				
MAG SPH	+	0.001***	0.016***	0.085***
MAG CYL	+	−0.004	−0.006	0.037
AGE	?	0.004***	0.008***	0.034***
GENDER (1 = male)	?	−0.036***	−0.064***	−0.322**
Personnel Decisions				
ln *EXPERIENCE*	−	−0.077***	−0.134***	−0.112***
Resource Allocation Decisions				
TRAINING	−	−0.063***	−0.044	−0.432
Avg *MAINT*	−	−0.019***	−0.045***	−0.149**
CENTER A (control)		1.085***	2.389***	7.948**
OVER (control)		−0.402***	−0.404***	−15.529
Adjusted R^2		23.69%	8.60%	
F-statistic		167.68***	51.51***	
Pseudo-R^2				18.12%
LR Statistic				393.60***

*,**,*** Indicates significance at the .05, .01, .001 levels, respectively (two-tailed tests).

[a] This table presents results of the estimation of the quality management function (n = 4,833). Models 1 and 2 are OLS estimations of *DEVIATION* and *DEVIATION*^{2}, respectively, modeled as a function of the quality-related decisions likely to be important in the current setting, and two control variables (an indicator for *CENTER A* and an indicator for overcorrections, *OVER*). Model 3 is a maximum likelihood (logistic) estimation of the *OUTTOL* quality measure as a function of the decision variables. Higher values of the *DEVIATION*, *DEVIATION*^{2}, and *OUTTOL* variables indicate lower quality. Higher values of *MAG SPH* and *MAG CYL* indicate higher service complexity. For ease of interpretation, all variables except the indicator variables are mean-centered.

Variable Definitions:

MAG SPH = the magnitude of the spherical equivalent correction needed (i.e., the severity of the myopia);
MAG CYL = the magnitude of the cylindrical correction needed (i.e., the severity of the astigmatism);
AGE = age of patient;
GENDER = gender of patient (1 = male);
ln *EXPERIENCE* = the log of the cumulative number of procedures performed by the surgeon;
TRAINING = indicator variable identifying the center-months following the firm-wide training program (1 = post-July 97);
Avg *MAINT* = mean maintenance expenditure per procedure for the center;
CENTER A = an indicator variable equal to 1 for Center A observations; and
OVER = an indicator variable equal to 1 for overcorrections.

Marketing Decisions

The results show significant relations between the various quality management choice variables and the forward-looking quality measure *DEVIATION*. Recalling that higher values of *DEVIATION* correspond to lower quality, the results show that quality is significantly associated with two patient characteristics: age and gender. Controlling for correction magnitudes, quality is decreasing (*DEVIATION* is increasing) in patient age (coefficient of

0.004, p-value < .001). Quality is also slightly better for men (coefficient of −0.036, p-value < .001) than for women. Finally, the quality measure is decreasing in service complexity as measured by the magnitude of the required spherical correction, *MAG SPH*. No statistically significant association is found between the quality measure and the magnitude of the cylindrical correction, *MAG CYL*.

The previous findings are potentially useful for marketing and advertising resource allocation decisions. For example, since quality is higher (and, hence, future warranty costs are lower) for men, relatively more resources can be directed toward advertising that appeals to men. Moreover, pricing can be adjusted to reflect the increased warranty costs associated with larger correction magnitudes. Indeed, firms in this industry have been moving toward discriminatory pricing where higher fees are charged for larger corrections.

Personnel Decisions

Consistent with prior research (e.g., Ittner et al. 2001), quality is increasing (*DEVIATION* is decreasing) in surgeon experience.[10] That is, more experienced surgeons achieve outcomes closer to *PLANO*. The coefficient on ln *EXPERIENCE* is −0.077 (p-value < .001). Thus, the experience effect is greatest for low values of *EXPERIENCE*—at the minimum *EXPERIENCE* of 11 procedures, every additional ten procedures improves the mean outcome by about 0.07 diopters—and diminishes at higher levels of experience (an improvement of 0.001 diopters for an additional ten procedures at the mean level of experience, and virtually no effect of experience at the maximum experience).[11]

Findings related to personnel decisions have direct implications for operational and management control. With evidence that experience improves quality, managers can focus on ways to move surgeons along the learning curve more quickly or to compensate for lack of experience. As an example, the current firm has developed technology and procedures to monitor outcomes for individual surgeons and to program into the equipment precise adjustments to compensate for differences arising from experience levels (e.g., differences in speed). Moreover, although firms cannot, by law, compensate physicians based on outcomes, the development of appropriately constructed forward-looking performance measures *can* be used for contracting in non-healthcare settings as a means of directing managerial attention to the long-term (Dikolli 2001).

Resource Allocation Decisions

Evidence that quality is increasing in investments in training and maintenance is also consistent with prior research. Specifically, deviations from the optimal outcome of *PLANO* are lower following the firm-wide training effort and for centers with higher per-procedure expenditures on maintenance. These results have direct uses for evaluating past investments and predicting the returns on future, similar ones. Moreover, such evaluations are made more timely with the use of contemporaneously determined forward-looking performance measures.

---

[10] Note that the model was estimated using the natural log of experience (measured as the cumulative number of procedures by the surgeon) to allow for diminishing returns. Although not the focus of this section, this illustrates another opportunity for management accountants, namely, to investigate ways to improve the measurement of the quality management function.

[11] That is, $\partial \mathit{DEVIATION} / \partial \mathit{EXPERIENCE} = -0.007 * (1/\mathit{EXPERIENCE})$. At the minimum *EXPERIENCE* level of 11 procedures, the improvement in *DEVIATION* associated with an additional ten procedures is $-0.077 * (1/11) * 10 = 0.07$ diopters. Similarly, at the mean *EXPERIENCE* level of 944 procedures, the improvement is $-0.077 * (1/944) * 10 = 0.001$ diopters. A *rough* rule of thumb is 1 diopter for every 100 (to 150) feet of visual acuity (e.g., 20/20 vs. 20/120).

Comparison of Alternative Measures

Models 2 and 3 in Table 7 present results of the estimation of the quality management function using the two alternative measures, *DEVIATION*^{2} and *OUTTOL*, respectively, as the quality outcome measure to be managed (i.e., the dependent variable). Inferences regarding quality management decisions are qualitatively similar as those for the quality outcome measure *DEVIATION*, with one important exception—namely, when using the measure shown to be a better predictor of future warranty costs (i.e., the more goal-congruent measure *DEVIATION*) as the quality outcome variable to be managed, improvements in quality following the training program are revealed. On the other hand, if managers in this setting were to use either of the other two quality measures proposed and evaluated in the previous section (i.e., *DEVIATION*^{2} or *OUTTOL*), then they would incorrectly infer that no future resources should be allocated to training since training appears to have no relation to the quality measure used. Thus, even with subtle measurement differences such as those that distinguish *DEVIATION* from the other two measures, important differences *can* arise in managerial decisions.

In summary, this illustration shows that the forward-looking quality measure is associated with quality management decisions likely to be important in the current setting. Thus, the estimation of a quality management function provides a tool for managers to evaluate past (and future) decisions based on their likely effect on quality and future warranty costs. The value of this tool lies in the measure of quality that is contemporaneous to the service delivery (and, therefore, more timely with respect to decisions) *and* has forward-looking properties.

VI. SUMMARY AND CONCLUSIONS

Prior research documents that nonfinancial performance measures are often leading indicators of financial performance. In this paper, I examine the effects of measurement alternatives on the forward-looking properties of a quality performance measure. Specifically, the quality and quality cost function measurement issues examined are: (1) the alternative *standards* against which quality is measured, (2) whether variations in quality have a *uniform* affect on future quality costs (i.e., consideration of an asymmetric relation between quality and future quality costs), and (3) the exact *functional form* or the relation between a quality measure and future quality costs. Research from operations, management, and marketing is used to develop hypotheses regarding the effect of these measurement alternatives on the strength of the forward-looking properties of the quality measure.

Using proprietary data from a medical services firm, I find, first, that the standard against which quality is measured affects the strength of the association between the quality measure and future quality-related warranty costs and, hence, the measure's role as a forward-looking performance indicator. Although marketing theory supports the use of the customer's *ex ante* expectation as the standard against which quality should be measured (e.g., Cronin and Taylor 1992, 1994; Parasuraman et al. 1985; Patterson 1993; Stank et al. 1999), an absolute, *optimal* outcome standard dominates an *expected* outcome standard in the current setting. Next, the warranty cost function is shown empirically to be asymmetric in the current setting. This result is consistent with the generalizable hypothesis that the strength of the relation between quality and future quality-related warrant costs is a function of customers' *expected net benefit* of the warranty work. Finally, alternative functional forms of the warranty cost function as suggested by current manufacturing practice and operations theory are considered. The empirical results do not support the tolerance limit approach that is commonly used in manufacturing settings and that assumes small variations from optimal quality are not costly for the firm. Instead, the results suggest that *all* deviations from optimal quality are costly in the current setting, consistent with quality theory in operations research (e.g., Taguchi et al. 1989). Moreover, the data support a linear (as opposed to an exponential) relation between the quality measure and future quality costs in the current setting.

To highlight the implications of this study for management accounting, the paper concludes with an illustration of how an "improved" forward-looking performance measure facilitates management decision making by capturing the effects of those decisions on future costs. For example, findings related to the association between the quality measure and expenditures on various conformance efforts can direct a firm to the optimal allocation of resources to those efforts (e.g., training seems to be more effective than the hiring of more skilled surgeons). Moreover, identifying the relation between various marketing decisions and the quality measure may suggest future advertising or pricing strategies. The fact that quality is measured contemporaneous to the service delivery *and* has forward-looking properties further strengthens its usefulness in evaluating past (and future) quality-related decisions in a timely manner. Finally, the illustration shows that decisions can differ when alternative forms of the quality measure are used.

As is characteristic of field research, the *specific* findings of this study are limited in their generalizability. However, the evidence provided is consistent with the very general proposition that the *construction* of a quality measure—that is, the selection of measurement alternatives—may improve its ability to predict future financial performance and, hence, its usefulness for decision making and control. Moreover, the measurement issues identified and the theories presented to suggest measurement alternatives arguably generalize beyond the measurement of quality and beyond the current setting. Finally, this paper speaks to a void in the management accounting research related to the factors that affect the strength of leading indicator relations. Additional theoretical and empirical identification of factors that affect leading indicator relations (e.g., customer satisfaction and future financial performance, Ittner and Larcker 1998a) remains an open and promising area for future research.

REFERENCES

Albright, T. L., and H. P. Roth. 1992. The measurement of quality costs: An alternative paradigm. *Accounting Horizons* 6 (2): 15–27.

Allison, P. D. 1995. *Survival Analysis Using the SAS System: A Practical Guide*. Cary, NC: SAS Institute Inc.

Anderson, E. W., and M. W. Sullivan. 1993. The antecedents and consequences of customer satisfaction for firms. *Marketing Science* 12 (2): 125–143.

———, C. Fornell, and D. R. Lehmann. 1994. Customer satisfaction, market share, and profitability: Findings from Sweden. *Journal of Marketing* 58 (3) 53–66.

Anderson, S. W. 1993. Measuring manufacturing flexibility: The impact of product mix complexity on operating performance and manufacturing overhead cost. Ph.D. dissertation, Harvard University.

———. 1995. Measuring the impact of product mix heterogeneity on manufacturing overhead cost. *The Accounting Review* 70 (3): 363–387.

Apte, U. M., U. S. Karmarkar, and R. Pitbladdo. 1997. Quality management in services: Analysis and measurement. In *The Practice of Quality Management*, edited by P. J. Lederer, and U. S. Karmarkar. Boston, MA: Kluwer Academic Publishers.

Balakrishnan, R., T. Gruca, and D. Nath. 1996. The effect of service capability on operating costs: An empirical study of Ontario hospitals. *Contemporary Accounting Research* 13 (1): 177–206.

Banker, R., and H. H. Johnston. 1993. An empirical study of cost drivers in the U.S. airline industry. *The Accounting Review* 68 (3): 576–601.

———, G. Potter, and D. Srinivasan. 2000. An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review* 75 (1): 65–92.

Bearden, W. O., and J. E. Teel. 1983. Selected determinants of consumer satisfaction and complaint reports. *Journal of Marketing Research* 20 (1): 21–28.

Cronin, J. J., and S. A. Taylor. 1992. Measuring service quality: A reexamination and extension. *Journal of Marketing* 56 (3): 55–68.

———, and ———. 1994. Modeling patient satisfaction and service quality. *Marketing Health Services* 14 (1): 34–44.

Crosby, P. B. 1979. *Quality is Free.* New York, NY: McGraw-Hill.

Datar, S. M., S. Kekre, T. Mukhopadhyay, and K. Srinivasan. 1993. Simultaneous estimation of cost drivers. *The Accounting Review* 68 (3): 602–614.

———, S. C. Kulp, and R. A. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39 (1): 75–92.

Darnell, A. C. 1994. *A Dictionary of Econometrics.* Lyme, NH: Edward Elgar Publishing, Inc.

Dikolli, S. S. 2001. Agent employment horizons and the contracting demand for leading indicators of future performance. *Journal of Accounting Research* 39 (3): 481–494.

Feigenbaum, A. V. 1983. Quality costs. In *Total Quality Control,* 3rd edition. New York, NY: McGraw-Hill.

Feltham, G. A., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review* 69 (3): 429–453.

Fine, C. H. 1986. Quality improvement and learning in productive systems. *Management Science* 32 (10): 1301–1315.

Foster, S. T., and E. E. Adam, Jr. 1996. Examining the impact of speed of quality improvement on quality-related costs. *Decision Sciences* 27 (4): 623–646.

Greene, W. H. 1997. Limited dependent variable and duration models. In *Econometric Analysis,* 3rd edition. Upper Saddle River, NJ: Prentice Hall, Inc.

Ireson, W. 1971. *Handbook of Industrial Engineering and Management.* 2nd edition. Englewood Cliffs, NJ: Prentice Hall Inc.

Ittner, C. D., and J. P. MacDuffie. 1995. Explaining plant-level differences in manufacturing overhead: Structural and executional cost drivers in the world auto industry. *Production and Operations Management* 4 (4): 312–334.

———. 1996. Exploratory evidence on the behavior of quality costs. *Operations Research* 44 (1): 114–130.

———, and D. Larcker. 1998a. Are non-financial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research* 36 (Supplement): 1–35.

———, and ———. 1998b. Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research* 10: 205–238.

———, V. Nagar, and M. V. Rajan. 2001. Empirical examination of quality learning models. *Management Science* 47 (4): 563–578.

Kim, M., and W. Liao. 1994. Estimating hidden quality costs with quality loss. *Accounting Horizons* 8 (1): 8–18.

Kinney, W. R., Jr. 2001. Accounting scholarship: What is uniquely ours? *The Accounting Review* 76 (2): 275–284.

Kotler, P., and P. N. Bloom. 1984. *Marketing Professional Services.* Englewood Cliffs, NJ: Prentice Hall.

Lambert, R. A. 1998. Customer satisfaction and future financial performance: Discussion of: Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research* 36 (Supplement): 37–46.

Love, C. E, R. Guo, and K. H. Irwin. 1995. Acceptable quality level versus zero-defects: Some empirical evidence. *Computers & Operations Research* 22 (4): 403–417.

MacArthur, J. B., and H. A. Stranahan. 1998. Cost driver analysis in hospitals: A simultaneous equations approach. *Journal of Management Accounting Research* 10: 280–312.

Nagar, V. 1998. The explanatory power of non-financial measures in bank earnings prediction. Working paper, University of Michigan, Ann Arbor, MI.

———, and M. Rajan. 2001. Revenue implications of product quality measures. *The Accounting Review* 76 (4): 495–513.

Nelson, E. C., R. T. Rust, A. Zahorik, and R. L. Rose. 1992. Do patient perceptions of quality relate to hospital financial results? *Journal of Health Care Marketing* 12 (4): 6–13.

Oliver, R. L. 1980. A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research* 17 (4): 460–469.

Parasuraman, A., V. Zeithaml, and L. Berry. 1985. A conceptual model of service quality and its implications for future research. *Journal of Marketing* (Fall): 41–50.

———, ———, and ———. 1994. Reassessment of expectations as a comparison standard in measuring service quality: Implications for further research. *Journal of Marketing* 58 (1): 111–124.

Patterson, P. G. 1993. Expectations and product performance as determinants of satisfaction for a high-involvement purchase. *Psychology & Marketing* 10 (5): 449–465.

Phillips, L. D., D. R. Chang, and R. Buzzell. 1983. Product quality, cost position, and business performance: A test of some key hypotheses. *Journal of Marketing* 47 (2): 26–43.

Stank, T. P., T. J. Vickery, and K. Shawnee. 1999. Effect of service supplier performance on satisfaction and loyalty of store managers in the fast food industry. *Journal of Operations Management* 17 (4): 429–447.

Taguchi, G., E. A. Elsayed, and T. Hsiang. 1989. *Quality Engineering in Production Systems*. NY: McGraw-Hill.

Zeithaml, V. A., L. L. Berry, and A. Parasuraman. 1996. The behavioral consequences of service quality. *Journal of Marketing* 60 (2): 31–46.

THE ACCOUNTING REVIEW
Vol. 78, No. 2
2003
pp. 581–604

Reviewers' Responses to Anticipated Stylization Attempts by Preparers of Audit Workpapers

Hun-Tong Tan
Nanyang Technological University

Ken T. Trotman
University of New South Wales

ABSTRACT: In this paper, we examine specific process gains and losses arising from the review process by considering the interaction between attributes of the preparer and attributes of the reviewer. Specifically, we investigate determinants of reviewers' performance in the presence of potential stylization attempts by audit workpaper preparers. We assess whether reviewers' detection of conclusion errors and documentation errors depends on three factors: nature of the preparer's stylization (by way of the preparer's differential emphasis on conclusion and documentation errors), auditor rank, and the reviewer's sensitivity to such stylization attempts. Our results show that, depending on the reviewers' stylization sensitivity, their cognitive templates can both facilitate and interfere with reviewers' effectiveness in responding to these stylization attempts, resulting in both potential gains and losses from the review process. Our paper identifies conditions related to whether and how reviewers can provide adaptive responses to stylization attempts, and demonstrates that conclusions about rank-related performance differences in the review process should consider complex interactions between attributes of the reviewer and the preparer. Implications are discussed.

Keywords: *audit review process; auditor rank; workpaper stylization.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

The audit review process, involving a reviewer (generally a more senior auditor) assessing the work quality of the preparer (generally a more junior auditor), is a major quality control mechanism designed to improve the quality of both workpaper documentation and conclusions made in the audit process. To evaluate the effectiveness of the review process, there is a need for greater understanding of the process gains and losses arising from the review process (Libby and Trotman 1993; Ramsay 1994).

We appreciate helpful comments from Mike Bamber, Terence Ng, Steve Salterio, and two anonymous referees. We thank the auditors from the Big 4 firms for participating in the study, Roger Gibson for assistance in developing the research instruments, and Mahreen Hasan and Amna Khalifa for research assistance. A research grant from the Australian Research Council is gratefully acknowledged.

Editor's note: This paper was accepted by E. Michael Bamber, *Ad Hoc* Editor.

Submitted April 2002
Accepted November 2002

Reviewers receive working papers intended to persuade them that the work performed, conclusions reached, and related documentation are appropriate and defensible. In order to persuade reviewers, workpaper preparers may engage in stylization—the manner with which preparers modify the audit approach and workpaper preparation in order to influence the reviewers' judgments and/or to conform to the preferences of the reviewers (Rich et al. 1997). A benefit of workpaper stylization is that it may lead to more efficient audits in that reviewers are more likely to accept workpapers that meet their specifications. However, there may be associated costs if stylization that conforms to the reviewers' preferences leads the reviewer to prematurely accept workpapers that contain erroneous documentation and/or conclusions.

In this study, we investigate reviewers' responses to potential stylization attempts by the preparer, and how the audit review process is influenced by an interaction between the attributes of the preparer and the attributes of the reviewer. We examine three variables: the manner with which a preparer is expected to stylize workpapers, the reviewer's sensitivity to such stylization attempts, and the reviewer's rank. How a preparer stylizes his/her workpaper preparation is a factor present in the audit environment that varies with the preparer and the reviewer with whom the preparer works with. While preparers can stylize their workpaper preparation in several ways to please the prospective reviewer, we examine a situation where the preparer is expected to stylize his/her workpaper preparation by focusing either on documentation or conclusion errors. We look at these aspects because, although such detection of documentation and conclusion errors are process gains arising from audit reviews (Ramsay 1994),[1] a preparer's stylization attempts can adversely affect reviewers' detection of such errors and the associated process gains associated with the review process.

We examine how two reviewer attributes interact in influencing the reviewer's response to these anticipated stylization attempts. One reviewer attribute relates to a reviewer's sensitivity to stylization, which, in our context, refers to the extent to which reviewers devote greater attention toward the detection of errors likely to be neglected by the preparer as a result of workpaper stylization. Reviewers who devote insufficient or inappropriate attention to neglected errors arising from stylization attempts likely compromise review effectiveness. The other reviewer attribute is the reviewer's knowledge, which refers to the domain knowledge and cognitive template a reviewer brings to bear on the task of reviewing the work of a preparer. In our study, we proxy reviewers' knowledge by their rank, as prior research suggests that audit managers have relatively stronger knowledge representations of conclusion (vs. documentation) errors, while the reverse is true of seniors (Ramsay 1994). This research suggests that during the review process, managers adopt conclusion error templates, while seniors adopt documentation error templates. We posit and demonstrate that auditors' responses to stylization attempts depend on auditor rank and their sensitivity to such stylization attempts.

We conduct an experiment in which managers and seniors review the workpapers of a preparer, who is known to stylize his/her work to please a reviewer and expected to stylize the workpaper preparation by paying more attention either to conclusion errors (and neglecting documentation errors), or to documentation errors (and neglecting conclusion errors). We find that when the preparer is expected to neglect documentation errors, seniors respond by detecting more documentation errors, but only when they are relatively more sensitive to stylization attempts. A different pattern of results is observed for managers—their detection of documentation errors first increases, then *falls* as their stylization sensitivity increases. In contrast, when the preparer is expected to neglect conclusion errors, managers are able to detect more conclusion errors with increasing stylization sensitivity. However, seniors' detection of conclusion errors is invariant across stylization sensitivity.

---

[1] The terms, documentation and conclusion errors, correspond to the terms mechanical and conceptual errors, respectively, used in Ramsay (1994). We use the terms documentation and conclusion errors in the paper and our experimental materials because pretests indicate that auditors are more familiar with these terms.

Our paper makes two primary contributions to the academic auditing literature. It provides greater understanding of the review process by examining reviewer behavior in response to preparer stylization. Prior research suggests that preparers stylize their workpapers (Rich et al. 1997; Gibbins and Trotman 2002), but there is mixed evidence on whether reviewers detect and correct for such stylization attempts (e.g., Yip-Ow and Tan 2000; Rich 2002). Our study suggests that two conditions must be met before reviewers can provide adaptive responses to preparers' stylization attempts. First, the reviewers must be sensitive to such attempts, and allocate appropriate attention toward the detection and correction of these stylization attempts. Second, the responses by the reviewers to these stylization attempts must match their review templates. Where their responses result in a mismatch between the task requirements and their review templates, performance may be adversely affected. On the other hand, where a match occurs, performance improves.

Our paper also contributes to the literature on auditor expertise related to rank differences, and on sources of performance differences between managers and seniors. In particular, our study adds to the literature on manager-senior differences in audit review performance (Ramsay 1994; Bamber and Ramsay 1997) by providing further evidence on how auditors' schemas (possessed by seniors and managers) can influence audit effectiveness. Existing literature suggests that audit managers outperform audit seniors in terms of detection of conclusion errors, but that the reverse is true of documentation errors (Ramsay 1994). However, these findings do not incorporate important features in the audit environment (e.g., that preparers stylize in different ways) and individual differences in sensitivity to such stylization attempts. We provide a more complete depiction of manager-senior differences by considering such complex interactions between individual attributes and the task environment.

The rest of the paper is organized as follows. We develop our hypotheses in Section II, and describe our research design in Section III. We present our results in Section IV, and conclude our paper in Section V.

II. HYPOTHESIS DEVELOPMENT

The review process is a control mechanism implemented by CPA firms to improve the quality of the workpaper documentation and conclusions made. The review process involves interactions between preparers of audit workpapers and their reviewers. Therefore, in as much as the reviewer assesses the preparer's work quality, the preparer can stylize his/her workpaper preparation and influence the reviewer's performance. To better understand reviewers' performance in the light of a preparer's stylization attempt, we posit a theoretical framework that involves three factors: the reviewer's underlying review template, the nature of the preparer's expected stylization attempt, and the reviewer's sensitivity to such a stylization attempt.

The first factor, reviewer's underlying review template, refers to the reviewer's cognitive template that he/she brings to bear on the task of reviewing a preparer's work. Arising from differences in responsibilities and resultant frequencies in processing different error types, reviewers have default review templates organized around the detection of different error types, specifically conclusion errors in the case of audit managers, or documentation errors in the case of audit seniors (Ramsay 1994); we term these the "default error types" for the purpose of this discussion. Ramsay's (1994) study indicates that managers' (seniors') default review templates enable them to detect, and possibly have a proclivity toward detecting, conclusion (documentation) errors. The use of these review templates leads to greater review efficiency and effectiveness in detecting default error types, which is particularly important in the context of time pressure and budget constraints. However, we argue that these default review templates can either facilitate or interfere with reviewers' effectiveness in responding to preparers' stylization attempts, depending on the context.

The second factor in our framework relates to a contextual factor—the nature of the preparer's expected stylization attempt, which refers to a particular form of stylization that is potentially used by the preparer (Rich et al. 1997). Prior research has investigated the effects of various stylization methods such as omitting evidence from audit workpapers (Ricchiute 1999), using justification memos that only highlight supporting evidence (Yip-Ow and Tan 2000), emphasizing evidence consistent and de-emphasizing evidence inconsistent with workpaper conclusions (Tan and Yip-Ow 2001), as well as exerting differential effort toward the task (Rich 2002). While there are different ways a preparer can stylize workpapers, we focus on a situation where a preparer, known to stylize his/her work to please a reviewer, is expected to focus on the detection of one error type—either conclusion errors or documentation errors (i.e., the *focused* error type)—at the expense of the other (i.e., the *neglected* error type).

Given this potential stylization by the preparer, the third factor in our framework—reviewer's sensitivity to preparer's stylization—becomes an important consideration. We define stylization sensitivity to be the extent to which a reviewer exerts greater attention toward detecting the error type not stylized for (i.e., the neglected error type) by the preparer.[2] Assessing an auditor's stylization sensitivity is important because directional attention toward detection of the neglected error type is, *ceteris paribus*, an appropriate response to increased risk of errors committed for this error type.[3]

These three factors interact to determine the extent to which a reviewer detects different kinds of errors in the presence of a preparer's differential attention to different error types. We posit that a reviewer's detection of different error types depends on the reviewer's stylization sensitivity, and whether his/her underlying review template matches the error type that is potentially neglected by the preparer. An overview of our theoretical framework is presented in Figure 1, Panel A.

Match between Neglected Error Type and Reviewers' Default Review Templates

For an error type potentially neglected by the preparer that matches the reviewer's default review template, increasing reviewer attention from greater stylization sensitivity toward detecting these errors should lead to correspondingly better detection of the neglected error type (see Figure 1, Panel A, Cell 1). Research in psychology is consistent with the view that such default templates, once activated (e.g., during the review process), prime people toward the default error types (Bower et al. 1979) and lead to superior performance in identifying these error types (e.g., Biederman et al. 1973; Shiffrin and Schneider 1977). In the context of the audit review process, therefore, a reviewer's well-developed default review template should facilitate the detection of template-congruent error

---

[2] Attention refers to the application of cognitive effort toward an activity, and the amount of attentional resources available to a decision maker at any point in time is assumed to be limited (Kahneman 1973). This constraint on attentional resources that a reviewer can devote to detecting a particular error type is exacerbated by time pressure ubiquitous in the audit environment (that is also present in our experiment). From this perspective, the reviewer's attention spent on detecting errors is better described as "constrained."

[3] This assumes that the costs of responding to the stylization attempt are not prohibitive and exceed the benefits.

FIGURE 1
Theoretical Framework and Hypotheses

Panel A: Overview of Theoretical Framework

	Reviewer's Default Review Template	
	Matches Error Type Neglected by Preparer	**Does Not Match Error Type Neglected by Preparer**
Predicted relationship between *neglected error type* detection and increasing stylization sensitivity	Increasing linear function as stylization sensitivity increases (Cell 1)	Inverted U-shaped function as stylization sensitivity increases (Cell 3)
Predicted relationship between *focused error type* detection and increasing stylization sensitivity	Null hypothesis (Cell 2)	Null hypothesis (Cell 4)

Our framework refers to a context where a workpaper preparer is expected to stylize workpaper preparation by focusing on one error type (i.e., the *focused* error type) and neglecting the other (i.e., the *neglected* error type). The reviewer's default review template refers to the cognitive template that he/she brings to bear on the task of reviewing a preparer's work, and this default template either matches or does not match the error type neglected by the preparer. Stylization sensitivity refers to the extent to which a reviewer exerts greater attention toward detecting the error type not stylized for (i.e., the neglected error type) by the preparer.

(*continued on next page*)

FIGURE 1 (continued)

Panel B: Summary of Hypotheses and Results

Auditor rank:	**Managers**				**Seniors**			
Preparer Focus:	**Doc** (*n* = 19)		**Conc** (*n* = 21)		**Doc** (*n* = 26)		**Conc** (*n* = 23)	
Nature of errors detected	**Doc**	**Conc**	**Doc**	**Conc**	**Doc**	**Conc**	**Doc**	**Conc**
Mean	7.68	8.47	7.62	9.43	7.38	8.46	7.91	9.00
(s.d.)	(1.80)	(1.26)	(2.04)	(1.66)	(1.81)	(1.21)	(1.90)	(1.68)
[range]	[4–11]	[6–11]	[3–12]	[6–12]	[4–11]	[6–11]	[5–11]	[6–12]
Stylization sensitivity								
Mean	0.26		0.30		0.42		0.35	
(s.d.)	(0.27)		(0.18)		(0.22)		(0.33)	
[range]	[–0.33 to 0.71]		[0.0 to 0.60]		[0.00 to 1.00]		[–0.80 to 0.71]	
Effect of Stylization Sensitivity (SS)	Null hypothesis	Positive	Inverted U shape	Null hypothesis	Null hypothesis	Inverted U shape	Positive	Null hypothesis
Hypothesis Number	H3a	H3b	H1a	H1b	H4a	H4b	H2a	H2b
Results supported	Yes	Yes	Yes	Yes	Partially	No	Yes	Yes

Doc = documentation errors; and
Conc = conclusion errors.

types (potentially neglected by the preparer) with increasing stylization sensitivity and greater vigilance.

Directional predictions of a reviewer's detection of the focused error type (i.e., error type on which the preparer is expected to focus) are more difficult because of counteracting forces that are potentially at play. There is a mismatch between this error type focused on by the preparer, and the reviewer's default review template, so performance on the detection of this error type is unlikely to be good to begin with (Nelson et al. 1995). Increased attention toward the neglected error type also implies correspondingly lower attention toward the complementary focused error type. On the other hand, prior research suggests that detection of one error type facilitates the detection of the complementary error type (Bamber and Ramsay 1997). Because these factors potentially offset each other, we do not make directional predictions on the reviewer's detection of the focused error type (see Figure 1, Panel A, Cell 2).

Mismatch between Neglected Error Type and Reviewers' Default Review Templates

When the preparer-neglected error type does not match the reviewer's default review template, we anticipate an inverted U-shaped relation between increasing stylization sensitivity and detection of the neglected error type (see Figure 1, Panel A, Cell 3). Consider a baseline condition where the reviewer is totally insensitive to stylization attempts by the preparer. In such a case, we expect the reviewer to apply his/her default review approach, and not pay additional attention toward detecting the neglected error type. In contrast, a reviewer sensitive to stylization attempts will shift his/her normal review approach toward detecting the neglected error type. At the initial stage, and within limits, this increased attention to the neglected error type should not create a significant mismatch with the reviewer's review template. As a result, at the initial stage, with increasing stylization sensitivity and greater attention toward the neglected error type, more of the neglected error type will be detected relative to the baseline condition.

However, the increased attention allocated toward detecting the neglected error type (that the reviewer does not, by default, focus on) changes the nature of the task that the reviewer is accustomed to, and increasingly exacerbates the mismatch between task characteristics and his/her default review template (Nelson et al. 1995). With increasing attention toward the neglected error, the mismatch between the task and default review template is magnified, interference occurs, and performance in detecting the neglected error type suffers (Bruner 1957). In psychology research, McDaniel and Kearney (1984) find that participants, instructed to use encoding strategies different from what they normally use, perform poorly relative to those using their default encoding strategies. Similarly, Singley and Anderson (1989) describe procedural interference, which occurs when the presence of a dominant skill interferes with the execution of another (see also Rehder 2001). In summary, we anticipate detection of the neglected error type to first increase, then decrease with increasing stylization sensitivity when the neglected error type is different from the default error type associated with the auditors' default review templates.

Again, it is difficult to make directional predictions of a reviewer's detection of the error type that the preparer is expected to focus on (i.e., the focused error type). Greater attention paid to the neglected error type can lead to lower attention paid to the focused error type, suggesting poorer detection of focused error type with increasing stylization sensitivity. However, this adverse effect may be mitigated. Psychology research (e.g., Biederman et al. 1973) and audit research (Nelson et al. 1995) suggest that performance is optimal when there is a matching of templates and task. These studies imply that because the focused error type matches the reviewer's default review template, detection of the focused error type is facilitated. Additionally, psychology research suggests that this matching results in more automatic processing with correspondingly lower demands on attentional resources (Shiffrin and Schneider 1977). As a result, detection of the focused error type can be sustained or less adversely affected, despite lower attention paid to the focused error type with increasing stylization sensitivity. Because the net effect might be that detection of the focused error type decreases or remains unaffected with increasing stylization sensitivity, we do not make directional predictions on the reviewer's detection of the focused error type (see Figure 1, Panel A, Cell 4).

In the subsections below, we apply the above framework to develop separate hypotheses for managers and seniors, who are expected to have different default review templates. A summary of the hypotheses is presented in Figure 1, Panel B.

Reviewers' Responses When Preparer Stylization Neglects Documentation Errors

Suppose a preparer is anticipated to stylize the workpapers by focusing on conclusion errors and neglecting documentation errors. We first discuss the responses of managers, who are expected to have a dominant conclusion error review template (Ramsay 1994). In this case, the neglected error type (documentation error) does not match the managers' default conclusion error review templates. Following the framework described above, managers' detection of documentation errors is expected to follow an inverted U-shaped function, and to first increase (relative to a baseline of zero stylization sensitivity) and then decrease with increasing stylization sensitivity.

With respect to conclusion errors (the error focused on by the preparer), it is conceivable that the mismatch can also adversely influence the detection of conclusion errors because the manager shifts attention toward the detection of documentation errors. A mitigating factor is that relative to documentation errors, there is a match between managers' default review template (relating to conclusion errors) and the error type focused by the preparer (also relating to conclusion errors). Because we do not have a basis to predict the overall effect, we posit a null hypothesis with respect to manager's detection of conclusion errors under these circumstances. Our hypotheses are formally expressed below:

H1a: For managers reviewing the work of a preparer who is expected to stylize workpapers by focusing on conclusion errors, detection of documentation errors will first increase and then decrease with increasing sensitivity to preparer stylization.

H1b: For managers reviewing the work of a preparer who is expected to stylize workpapers by focusing on conclusion errors, detection of conclusion errors will be unaffected by their sensitivity to preparer stylization.

Seniors possess dominant templates for the detection of documentation errors (Ramsay 1994). Therefore, in the situation where the preparer is expected to focus on detecting conclusion errors (and neglect the detection of documentation errors), there is a match between seniors' dominant templates for documentation error detection and the neglected error type. Our framework predicts that the seniors' relatively well-developed documentation error templates should facilitate the detection of documentation errors with increasing stylization sensitivity. With respect to seniors' detection of conclusion errors, increasing attention paid to documentation error detection may lead to less attentional resources devoted to, and poorer performance in, conclusion errors detection. However, detection of documentation errors might also play a facilitative role in the detection of conclusion errors (Bamber and Ramsay 1997). Because the overall effect can go either way, we do not make a directional hypothesis on seniors' detection of conclusion errors here.

H2a: For seniors reviewing the work of a preparer who is expected to stylize workpapers by focusing on conclusion errors, detection of documentation errors will increase with increasing sensitivity to preparer stylization.

H2b: For seniors reviewing the work of a preparer who is expected to stylize workpapers by focusing on conclusion errors, detection of conclusion errors will be unaffected by their sensitivity to preparer stylization.

Reviewers' Responses When Preparer Stylization Neglects Conclusion Errors

Consider, next, a case where the preparer is anticipated to stylize the workpapers by focusing on documentation errors (and neglecting the detection of conclusion errors). In this case, managers' dominant conclusion error detection templates match the neglected error type (conclusion error), and we expect more conclusion errors to be detected with greater vigilance associated with increasing stylization sensitivity. With respect to the detection of the focused error type (documentation errors), as in the earlier situations, counteracting forces are at play, and we, therefore, express this prediction as a null hypothesis.

H3a: For managers reviewing the work of a preparer who is expected to stylize workpapers by focusing on documentation errors, detection of documentation errors will be unaffected by their sensitivity to preparer stylization.

H3b: For managers reviewing the work of a preparer who is expected to stylize workpapers by focusing on documentation errors, detection of conclusion errors will increase with increasing sensitivity to preparer stylization.

Audit seniors are expected to have review templates where the detection of documentation errors (vs. conclusion errors) dominates. When the preparer stylizes by focusing on detecting documentation errors, there is a mismatch between the neglected error type (conclusion error) and the seniors' dominant templates for documentation error detection. Following our framework, we posit that seniors' detection of conclusion errors will follow an inverted U-shaped function with increasing stylization sensitivity. As with our earlier discussion, we express our prediction related to the focused error type (documentation errors) as a null hypothesis.

H4a: For seniors reviewing the work of a preparer who is expected to stylize workpapers by focusing on documentation errors, detection of documentation errors will be unaffected by increasing sensitivity to preparer stylization.

H4b: For seniors reviewing the work of a preparer who is expected to stylize workpapers by focusing on documentation errors, detection of conclusion errors will first increase and then decrease with increasing sensitivity to preparer stylization.

III. METHOD

Participants

We recruited our participants from the Sydney and Melbourne offices of one Big 4 accounting firm, by having the staff training manager send a memo (endorsed by a senior partner) requesting research volunteers, with the promise of A$75 (U.S.$38) as an incentive. Altogether, 49 audit seniors (mean experience = 42.8 months) and 40 audit managers (mean experience = 91.2 months) participated in the study. These participants completed the research instrument in their firm's offices in the presence of one of the authors. On average, the task took about 70 minutes to complete.

Design

Our study involves three independent variables: *Preparer Focus*, *Rank*, and *Stylization Sensitivity*. The first independent variable, *Preparer Focus*, is a between-subject variable manipulated at two levels. In the *Preparer Focus (documentation error)* condition, the preparer is known to stylize his workpapers and is expected to focus on documentation errors and neglect conclusion errors. In the *Preparer Focus (conclusion error)* condition, the preparer is known to stylize his workpapers and is expected to focus on conclusion errors and neglect documentation errors. The second variable *Rank* refers to audit managers and audit seniors. Finally, the variable *Stylization Sensitivity* is measured using scenarios (described below), presented at the end of the experiment, that measure participants' differential effort levels in response to different stylization attempts. We measured two dependent variables: the number of conclusion errors and the number of documentation errors correctly identified in a recognition test.

Procedure and Research Materials

At the start of the experiment, we provided participants with separate envelopes containing the instrument, along with an introductory package, and instructed participants to open one envelope at a time. The introductory package contained the background information. We informed participants that the study required them to review the inventory section of an audit file for a hypothetical client. They were told to assume that they had just been assigned to take over the manager's role in the audit of an existing client.[4] In all conditions, participants were informed that the audit workpapers for this client had not been reviewed but needed to be reviewed urgently because of the tight deadline. However, the previous in-charge manager, Kate Mitchell, had suddenly taken ill and was therefore unavailable to conduct the review. Following this information, participants read the background of the preparer (Paul Smith), who had worked with the previous manager-in-charge on three occasions. Participants learned that the preparer was known to first find out the reviewer's style and approach to workpaper review, and then to gear his workpapers toward satisfying the reviewer. Specifically, the preparer was known to pay relatively more attention to identifying conclusion or documentation errors depending on the preferences of the manager. Information about Kate (the previous manager) was then varied, depending on the experimental condition. Participants in the *Preparer Focus (conclusion error)* condition learned that Kate was known to be particularly concerned about the appropriateness of approach and conclusion, and that some partners had commented that, although the working papers that had gone through her review contained very few conclusion errors, the working papers periodically contained documentation errors. Participants in the *Preparer Focus (documentation error)* condition learned that Kate was known to be particularly concerned about documentation errors and that some partners had commented that, although the working papers that had gone through her review contained very few documentation errors, the working papers periodically contained conclusion errors.

---

[4] We kept constant the rank of the reviewer that participants adopted to avoid introducing differences in participants' perceptions of the types of errors that auditors at a particular rank are *expected* to detect. By having a context where seniors took over a manager's duties, we reduced the likelihood that our results for seniors were obtained because our participating seniors believed that they *had to* play the role of seniors and had to adopt documentation review templates (associated with seniors' review responsibilities). Holding constant the instructed role of the reviewer provides a more powerful test of our theory, and also biases against our finding rank-related differences in participants' responses.

To ensure that participants reflected upon an approach to reviewing the workpapers, we imposed time pressure on the participants at this stage by telling them that, although other participants had taken an hour to complete the review, they would be given only 30 minutes to work on the review task. We determined the time limit of 30 minutes based on pretests. The presence of time pressure suggested to participants that they would not have time to spend on all aspects of the review and implied the need to consider a review approach in light of this time pressure.[5] Participants were also asked to think about how they would approach the review of the working papers in the light of the above information and to describe their proposed audit approach.

Following this, participants opened Envelope A that contained the working paper prepared by Paul Smith. The workpapers included information on the company background, trial balances, schedule of adjusting entries, and work performed by a subordinate on various audit objectives pertaining to the inventory section. Participants were instructed to complete the task within 30 minutes, and were allowed to take notes.

They then opened Envelope B, and answered 24 true/false questions relating to the workpapers. We designed the working papers so that there were 12 documentation errors and 12 conclusion errors, resulting in a total of 24 items for the recognition test. For each error type, half were actual errors found in the workpapers, and half were "false" errors in that the particular items were properly prepared in the workpapers.[6] Participants' accuracy in answering these questions formed our dependent measures.[7] We created two versions of these test questions such that answers were true in one version and false in the other version, and vice versa. Participants were randomly allocated to these versions and all treatments. After completing the materials in Envelope B, participants opened the remaining envelopes, one at a time, which contained questions related to manipulation checks, stylization sensitivity, and knowledge measures.

To measure our third independent variable, *Stylization Sensitivity*, we presented participants with two scenarios that were modeled after the settings in the *Preparer Focus* condition. Participants were told to assume that they had been assigned to take over the manager's role in the audit of a client because the previous manager-in-charge had taken ill. The senior on the job had previously worked with that manager on earlier assignments, and was known to stylize working papers based on the preferences of the reviewer. Participants were instructed to indicate the percentage of effort they would allocate to detect documentation and conclusion errors in these scenarios. In Scenario 1 (Scenario 2), the ill manager was particularly concerned about documentation (conclusion) errors and was significantly more tolerant of conclusion (documentation) errors. The differential and directional effort that participants allocated to documentation vs. conclusion errors, for Scenarios 1 and 2, is a measure of auditors' sensitivity in responding to stylization attempts.

---

[5] If participants experienced no time pressure, then they would be less likely differentially to pay attention to different error types depending on preparer focus (and may pay attention to all error types since there was no time pressure). This would have been uncharacteristic of actual practice where time pressure is common.

[6] After running the experiment, we discovered that one of our actual documentation errors was a "false" error. As a result, we ended up with 12 documentation errors (5 actual, 7 "false") and 12 conclusion errors (6 actual, 6 "false").

[7] To assess whether the dependent measures are influenced by guessing among participants, we compute detectability indices d' (Tanner and Swets 1954) for both documentation and conclusion errors detected. These indices adjust for "false-alarm" rates (e.g., the tendency to respond "true" even when the answer is "false"). We obtain results similar to our main analyses.

IV. RESULTS

Manipulation Checks

To assess our *Preparer Focus* manipulation, we asked participants whether the ill manager, Kate Mitchell, was known to be particularly concerned about conclusion or documentation errors. All participants in the *Preparer Focus (conclusion error)* and *Preparer Focus (documentation error)* conditions correctly responded that the ill manager was concerned about conclusion and documentation errors, respectively. We also asked participants several questions about the preparer, Paul Smith, using 11-point scales (e.g., 0 = extremely unlikely to fit the preferences of reviewer; 10 = extremely likely to fit preferences of reviewer). Participants provided a mean rating of 9.53 on the extent to which Paul was likely to prepare workpapers to fit the preferences of the reviewer. Compared to participants in the *Preparer Focus (conclusion error)* condition, those in the *Preparer Focus (documentation error)* condition felt that Paul was more likely to focus on conclusion errors (means = 2.28 and 9.18, respectively; p = .000), and less likely to focus on documentation errors (means = 9.34 and 1.56, p = .000). These findings suggest a successful *Preparer Focus* manipulation.

Stylization Sensitivity

We measure *Stylization Sensitivity* by comparing the effort participants (expressed in percentage terms) indicate they would allocate to detecting documentation vs. conclusion errors under Scenario 1 (preparer focuses on documentation errors) and Scenario 2 (preparer focuses on conclusion errors). We use this operationalization of *Stylization Sensitivity* because alternative measures involving directly asking participants about their sensitivity to stylization attempts likely involve significant noise.

Using the percentage of effort allocated to detecting documentation errors as a unit of analysis, we compute a ratio, *Stylization Sensitivity*, defined as follows:

$$\frac{\{\text{Documentation errors effort (Scenario 2)} - \text{Documentation errors effort (Scenario 1)}\}}{\{\text{Documentation errors effort (Scenario 2)} + \text{Documentation errors effort (Scenario 1)}\}}$$

Note that in Scenario 2 (Scenario 1), the preparer is likely to neglect documentation (conclusion) errors. Reviewers who are sensitive to the preparer's stylization attempt, specifically the errors potentially neglected by the preparer, would indicate exerting *more* effort toward detecting documentation errors in Scenario 2, and *less* effort toward detecting documentation errors (which corresponds to more effort toward detecting conclusion errors) in Scenario 1. A higher *Stylization Sensitivity* ratio therefore indicates greater sensitivity toward detecting error types likely neglected as a result of the preparer's stylization attempts.

The *Stylization Sensitivity* ratio can be positive, zero, or negative. A positive ratio suggests that the reviewer is sensitive to neglected errors arising from the preparer's stylization attempt, where a ratio of 1 indicates extremely high sensitivity, involving a participant allocating 100 percent effort to detecting documentation errors in Scenario 2, and 0 percent effort to detecting documentation errors in Scenario 1. A ratio of 0 suggests that the reviewer is insensitive to stylization attempts, allocating equal effort to the detection of documentation errors regardless of the nature of preparer's stylization attempt. A negative ratio suggests that the reviewer may be responding dysfunctionally, allocating lower effort to detecting documentation errors in Scenario 2 (when greater effort should have been allocated), and greater effort to detecting documentation errors in Scenario 1 (when lower effort should have been allocated).

Figure 1, panel B shows the descriptive statistics for the *Stylization Sensitivity* ratio, partitioned by *Rank* and *Preparer Focus*. There is variation in the ratios within each condition, with ratios ranging from –0.80 to 1.00. Eight out of the 89 participants have a ratio of 0, and only three (two seniors, one manager) have negative ratios.[8]

Participants' responses to debriefing questions provide insights into their observations of stylization in practice. Participants reported that, among the audits that they had conducted, they observed stylization by their subordinates in terms of procedures to be performed (in 37 percent of all audits), conclusions made (in 32 percent of all audits), and documentation approach (45 percent of all audits). On a scale of 1 (not at all) to 9 (all the time), auditors recorded a mean rating of 3.89 in terms of the extent they specifically looked out for preparer stylization in their audit engagements.[9] These descriptive statistics indicate that auditors do observe stylization among preparers, that stylization occurs in a fairly large proportion of audits, and that stylization occurs through different means (including the stylization of documentation and conclusion errors that we examine in this study).

Main Analyses

Our hypotheses suggest that depending on auditor rank and error type, there will be inverted U-shaped relations between errors detected and *Stylization Sensitivity* (henceforth, *SS*) under some conditions (H1a and H4b), and linear relations under others (H2a and H3b). We follow Aiken and West's (1991) procedure for testing curvilinear relationships, and conduct separate regression analyses for managers and seniors. We mean-center the *SS* variable within rank to avoid collinearity problems (Aiken and West 1991).[10] The dependent variables are documentation errors detected and conclusion errors detected, and independent variables are *Preparer Focus* (coded as 0 = documentation error focus, 1 = conclusion error focus), *SS*, the quadratic term of *SS* (SS^2), the *Preparer Focus* by *SS* interaction, and the *Preparer Focus* by SS^2 interaction.[11] The formal regression equation is in the form:

$$\text{Errors detected} = b_0 + b_1 SS + b_2 SS^2 + b_3 PF + b_4 SS \times PF + b_5 SS^2 \times PF. \quad (1)$$

The quadratic term SS^2 relates to the curvilinear aspect of the relationship. For a simple equation that contains only *SS* and SS^2 terms (that we analyze later), an inverted U-shaped curve implies that the coefficient associated with SS^2 should be significantly negative, and the coefficient associated with *SS* should not be significantly different from zero (Aiken and West 1991).[12] With the presence of another independent variable and the existence of interaction terms, as in Equation (1) above, interpretation becomes more complex and the variables in Equation (1) can be rearranged as follows:

---

[8] We obtain similar results when we delete these three observations.

[9] We also assess whether there are rank differences in these responses. We find no rank differences ($p > .213$) for three of the four questions, the exception being the question on stylization of procedures, where percentages reported are significantly higher for seniors than managers (means = 45.3 and 29.7, respectively; $p = .008$).

[10] The *SS* variable does not vary with either *Rank*, *Preparer Focus*, or the interaction of the two variables ($p > .438$).

[11] Our hypotheses suggest a four-way interaction among *Auditor Rank*, *Preparer Focus*, SS^2, and *Error Type* (documentation errors detected, conclusion errors detected). As an overall test of these hypotheses, we re-perform this regression (also described in Table 1) by adding as additional independent variables, *Auditor Rank*, *Error Type* (as a repeated measure), the two-way interaction between these variables, and all possible interactions involving these variables and other variables (*Preparer Focus*, *SS*, and SS^2). Consistent with our hypotheses, the four-way interaction among *Auditor Rank*, *Preparer Focus*, SS^2, and *Error Type* is statistically significant ($p = .018$).

[12] Both linear (*SS*) and quadratic (SS^2) terms are to be included in the regression equation, even though only the quadratic term is relevant for assessing the existence of a curvilinear relationship (Aiken and West 1991, 63). The linear term controls for the overall linear relationship in the data.

$$\text{Errors detected} = (b_0 + b_3PF) + (b_1 + b_4PF)SS + (b_2 + b_5PF)SS^2. \tag{2}$$

The $(b_2 + b_5PF)$ coefficient for SS^2 indicates the nature of the curvilinear relation between errors detected and *SS* at specific values of *Preparer Focus*. Since it is not possible to ascertain the statistical significance of the $(b_2 + b_5PF)$ coefficient from Equation (1), a statistically significant *Preparer Focus* by SS^2 interaction term in Equation (1) only allows us to draw the general interpretation that the nature of the curvilinear relation between errors detected and *SS* varies by *Preparer Focus*. Specific tests of the nature and direction of slopes are reported in Tables 2 and 3.

Table 1, Panels A and B, report the results of the separate regressions for managers and seniors, respectively. With documentation errors detected as a dependent variable, significant SS^2 by *Preparer Focus* interactions are obtained for both managers ($p = .010$) and seniors ($p = .004$). With conclusion errors detected as a dependent variable, the SS^2 by *Preparer Focus* interaction is marginally significant for managers ($p = .080$), but not statistically significant for seniors ($p = .530$). These results indicate that the magnitude and nature of the curvilinear interactions vary by *Preparer Focus* and rank.[13]

As specific tests of our hypotheses, we conduct separate regressions for each level of *Rank* and *Preparer Focus*, with the linear and quadratic dimensions of *SS* as independent variables, and test for simple slopes at different levels of *SS*. Results of the regressions and tests of simple slopes are shown in Tables 2 and 3, respectively.[14]

In H1a, we predict that managers' detection of documentation errors will first increase, then decrease with increasing *SS* sensitivity. As indicated in Table 2, Panel A, for managers' detection of documentation errors in the *Preparer Focus (conclusion error)* condition, consistent with an inverted U-shaped curve (Aiken and West 1991), the quadratic function of *SS* is negative and significant ($p = .016$), while the linear function of *SS* is statistically insignificant from zero ($p = .266$); one-tailed tests. As more direct tests of our hypotheses, we test the nature of the simple slopes (change in documentation errors detected per unit change in *SS*; Aiken and West 1991) at three points—at the mean value of *SS*, and 1 standard deviation of *SS* above and below its mean. Consistent with our expectation, Table 3, Panel A shows that the slope is significantly positive 1 standard deviation below the mean ($p = .047$; one-tailed test), not significant at the mean ($p = .266$), and significantly

---

[13] Residual analyses indicate that assumptions of linearity and homogeneity of variance are met. The normal probability plots also indicate normality of the error terms. None of the independent variables has variance inflation factor (VIF) exceeding 10, the suggested cut-off point indicative of multicollinearity (see Neter et al. 1996), except for the SS^2 term (VIF = 20.89) and the SS^2 × *Preparer Focus* term (VIF = 20.34). However, after the deletion of three outliers (based on Cook's distance), there is no evidence of multicollinearity as the VIF for these two terms fall below 6.13 (VIF for the other terms < 7.91). Deletion of the outliers produce results similar to our full sample, and we report our results using the full sample.

[14] We control for auditors' experience within rank (*Rank experience*) and experience in a manufacturing sector (*Manufacturing experience*), and find similar results with or without these covariates.

TABLE 1
Regression Results of Effects of Stylization Sensitivity and Preparer Focus on Documentation Errors and Conclusion Errors Detected, Segregated by Rank

Panel A: Audit Managers Sample ($n = 40$): Preparer Focus and Stylization Sensitivity as Independent Variables

	Documentation Errors Detected as Dependent Variable				Conclusion Errors Detected as Dependent Variable			
	Regression Coefficient	Std. Error	t	p-value	Regression Coefficient	Std. Error	t	p-value
Intercept	7.225	.536	13.487	.000	8.389	.447	18.766	.000
PF	1.396	.765	1.826	.077	1.593	.638	2.497	.018
SS	.022	.365	.060	.953	.163	.305	.534	.597
SS^2	.333	.255	1.303	.201	.074	.213	.345	.732
$PF \times SS$	−.441	.643	−.685	.498	.262	.536	.488	.629
$PF \times SS^2$	−1.935	.703	−2.753	.010	−1.061	.586	−1.809	.080
	$R^2 = 0.232$, Adjusted $R^2 = 0.119$				$R^2 = 0.188$, Adjusted $R^2 = 0.069$			

Panel B: Audit Seniors Sample ($n = 49$): Preparer Focus and Stylization Sensitivity as Independent Variables

	Documentation Errors Detected as Dependent Variable				Conclusion Errors Detected as Dependent Variable			
	Regression Coefficient	Std. Error	t	p-value	Regression Coefficient	Std. Error	t	p-value
Intercept	7.826	.397	19.690	.000	8.264	.345	23.967	.000
PF	−.377	.559	−.674	.504	.583	.485	1.201	.236
SS	.807	.500	1.616	.113	−.292	.433	−.674	.504
SS^2	−.850	.386	−2.202	.033	.369	.335	1.100	.277
$PF \times SS$	.286	.689	.415	.680	.627	.597	1.049	.300
$PF \times SS^2$	1.280	.415	3.083	.004	−.229	.360	−.635	.530
	$R^2 = 0.250$, Adjusted $R^2 = 0.162$				$R^2 = 0.084$, Adjusted $R^2 = -0.022$			

PF = preparer focus (0 = documentation error focus, 1 = conclusion error focus);
SS = stylization sensitivity; and
SS^2 = the quadratic term of SS.

negative 1 standard deviation above the mean ($p = .008$; one-tailed test).[15] Overall, these results are consistent with H1a.

Hypothesis 1b is a null hypothesis relating to the effect of stylization sensitivity on managers' detection of conclusion errors in the *Preparer Focus (conclusion error)* condition. We find no statistically significant linear or quadratic functions of stylization sensitivity ($p = .876$ and $.247$, respectively). Tests of simple slopes indicate that the slopes are not

---

[15] Consistent with an inverted U-shaped curve, the regression equation (in Table 1) yields predicted values of 7.44, 8.62, and 6.60 at 1 standard deviation below, at, and 1 standard deviation above the mean of *SS*, respectively. These numbers also provide some evidence on the magnitude of the performance gains and losses from increasing stylization in this instance. As stylization sensitivity increases (from 1 standard deviation below to the mean of *SS*), the performance increase is 1.18 errors (equivalent to 15.9 percent out of a base of 7.44, or 9.8 percent out of the maximum 12 errors possibly detectable). From the mean of *SS* to 1 standard deviation above *SS*, the performance decrement is 2.02 (equivalent to 23.4 percent out of a base of 8.62, or 16.8 percent out of the maximum possible errors detectable).

TABLE 2
Regression Results of Effects of Stylization Sensitivity (*SS*) on Documentation Errors and Conclusion Errors Detected, Segregated by Preparer Focus and Rank

Panel A: Preparer Focus (Conclusion Error), Audit Managers Sample (n = 21)

	Documentation Errors Detected as Dependent Variable [H1a]				Conclusion Errors Detected as Dependent Variable [H1b]			
	Magnitude of Simple Slope	Std. Error	t	p-value	Magnitude of Simple Slope	Std. Error	t	p-value
SS	−.298	.467	−.639	.266*	.074	.468	.158	.876
SS^2	−1.356	.578	−2.347	.016*	−.694	.579	−1.198	.247

Panel B: Preparer Focus (Conclusion Error), Audit Seniors Sample (n = 23)

	Documentation Errors Detected as Dependent Variable [H2a]				Conclusion Errors Detected as Dependent Variable [H2b]			
	Magnitude of Simple Slope	Std. Error	t	p-value	Magnitude of Simple Slope	Std. Error	t	p-value
SS	.743	.430	1.728	.050*	−.117	.437	−.267	.792
SS^2	.344	.139	2.477	.011*	.050	.141	.352	.728

Panel C: Preparer Focus (Documentation Error), Audit Managers Sample (n = 19)

	Documentation Errors Detected as Dependent Variable [H3a]				Conclusion Errors Detected as Dependent Variable [H3b]			
	Magnitude of Simple Slope	Std. Error	t	p-value	Magnitude of Simple Slope	Std. Error	t	p-value
SS	.024	.367	.065	.949	.270	.188	1.435	.086*
SS^2	.301	.257	1.171	.259	.140	.132	1.061	.152*

Panel D: Preparer Focus (Documentation Error), Audit Seniors Sample (n = 26)

	Documentation Errors Detected as Dependent Variable [H4a]				Conclusion Errors Detected as Dependent Variable [H4b]			
	Magnitude of Simple Slope	Std. Error	t	p-value	Magnitude of Simple Slope	Std. Error	t	p-value
SS	.820	.495	1.655	.111	−.333	.355	−.939	.179*
SS^2	−.834	.383	−2.180	.040	.356	.274	1.299	.104*

* One-tailed test.

significantly different from zero either 1 standard deviation below the mean, at the mean, or 1 standard deviation above the mean of *SS* ($p > .255$).

Hypotheses 2a and 2b relate to seniors' detection of documentation and conclusion errors, respectively, in the *Preparer Focus (conclusion error)* condition. Specifically, in H2a, we anticipate that seniors' detection of documentation errors will increase with increasing

TABLE 3
Test of Simple Slopes at Different Levels of Stylization Sensitivity (*SS*), with Documentation Errors and Conclusion Errors Detected as Dependent Variables

Panel A: Preparer Focus (Conclusion Error), Audit Managers Sample (n = 21)

	Documentation Errors Detected as Dependent Variable [H1a]				**Conclusion Errors Detected as Dependent Variable [H1b]**			
	Magnitude of Simple Slope	**Std. Error**	**t**	**p-value**	**Magnitude of Simple Slope**	**Std. Error**	**t**	**p-value**
Simple slope at 1 standard deviation below mean of *SS*	2.414	1.366	1.767	.047*	1.461	1.369	1.067	.300
Simple slope at mean of *SS*	−0.298	0.467	−.639	.266*	0.074	.468	.158	.876
Simple slope at 1 standard deviation above mean of *SS*	−3.000	1.114	−2.703	.008*	−1.313	1.116	−1.177	.255

Panel B: Preparer Focus (Conclusion Error), Audit Seniors Sample (n = 23)

	Documentation Errors Detected as Dependent Variable [H2a]				**Conclusion Errors Detected as Dependent Variable [H2b]**			
	Magnitude of Simple Slope	**Std. Error**	**t**	**p-value**	**Magnitude of Simple Slope**	**Std. Error**	**t**	**p-value**
Simple slope at 1 standard deviation below mean of *SS*	.056	.280	.200	.423*	−0.216	.284	−.760	.456
Simple slope at mean of *SS*	.743	.430	1.728	.050*	−0.117	.437	−.267	.792
Simple slope at 1 standard deviation above mean of *SS*	1.431	0.668	2.144	.022*	−0.017	.678	−.026	.980

Panel C: Preparer Focus (Documentation Error), Audit Managers Sample (n = 19)

	Documentation Errors Detected as Dependent Variable [H3a]				**Conclusion Errors Detected as Dependent Variable [H3b]**			
	Magnitude of Simple Slope	**Std. Error**	**t**	**p-value**	**Magnitude of Simple Slope**	**Std. Error**	**t**	**p-value**
Simple slope at 1 standard deviation below mean of *SS*	−.577	.536	−1.077	.298	−.009	.275	−.034	.486*
Simple slope at mean of *SS*	.024	.367	.065	.949	.270	.188	1.435	.086*
Simple slope at 1 standard deviation above mean of *SS*	.625	.714	.876	.394	.550	.366	1.503	.077*

(*continued on next page*)

TABLE 3 (continued)

Panel D: Preparer Focus (Documentation Error), Audit Seniors Sample (n = 26)

	Documentation Errors Detected as Dependent Variable [H4a]				**Conclusion Errors Detected as Dependent Variable [H4b]**			
	Magnitude of Simple Slope	**Std. Error**	**t**	**p-value**	**Magnitude of Simple Slope**	**Std. Error**	**t**	**p-value**
Simple slope at 1 standard deviation below mean of *SS*	2.484	1.105	2.248	.035	−1.044	.793	1.312	.101*
Simple slope at mean of *SS*	.820	.495	1.655	.111	−.333	.355	−.939	.179*
Simple slope at 1 standard deviation above mean of *SS*	−0.844	.659	−1.281	.213	.378	.472	.801	.216*

* One-tailed test.

sensitivity to preparer stylization. Results in Table 2, Panel B, shows that both the linear and quadratic functions of *SS* are positive and statistically significant (p = .050 and .011, respectively; one-tailed test), which suggests an upward-sloping curve (Aiken and West 1991).[16] Tests of the simple slopes show that the slope is statistically insignificant from zero at 1 standard deviation below the mean of *SS*, but significantly positive at the mean of *SS* and 1 standard deviation above the mean (p = .423, .050, and .022, respectively; one-tailed tests). This result is consistent with our prediction in H2a.

We do not reject the null hypothesis posited in H2b as we find that neither the linear nor quadratic function of *SS* is significantly associated with seniors' detection of conclusion errors in the *Preparer Focus (conclusion error)* condition (p = .792 and .728, respectively). Tests of simple slopes at 1 standard deviation below, at, and above the mean of *SS* also reveal that these slopes are not significantly different from zero (p > .456).

The set of hypotheses, H3a and H3b, relates to managers' detection of different error types in the *Preparer Focus (documentation error)* condition. Hypothesis 3a specifies that increasing *SS* sensitivity will not influence managers' detection of documentation errors, while H3b predicts that *SS* sensitivity will positively impact their detection of conclusion errors. Consistent with H3a, we find no statistically significant linear (p = .949) or quadratic effect (p = .259) of *SS* on managers' detection of documentation errors, see Table 2, Panel C. Tests of simple slopes at different levels of *SS* indicate that the slopes are not significantly different from zero (p > .298), see Table 3, Panel C. With respect to managers' detection of conclusion errors, we find a marginally significant positive linear effect of *SS* (p = .086; one-tailed test), but not a significant quadratic *SS* effect (p = .152; one-tailed test). Tests of simple slopes show no statistically significant effect 1 standard deviation below the mean of *SS* (p = .486; one-tailed test), but marginally positive effects at the mean and 1 standard deviation above the mean of *SS* (p = .086 and .077, respectively; one-tailed tests). These results support our prediction in H3b.

---

[16] The positive and significant coefficients for both *SS* and SS^2 suggest that errors detected increases across all values of *SS*, but not proportionately so. Specifically, the simple slopes should be increasingly more positive as *SS* increases (Aiken and West 1991), a pattern supported by the test of simple slopes.

Hypothesis 4a states that increasing *SS* sensitivity will not influence seniors' detection of documentation errors in the *Preparer Focus (documentation error)* condition. Results in Table 2, Panel D show that the linear function is positive (p = .111) and quadratic function of *SS* is negative (p = .040). Contrary to the null hypothesis, tests of simple slopes shown in Table 3, Panel D reveal a statistically significant positive slope 1 standard deviation below the mean of *SS* (p = .035), although the slopes are not statistically significant at the mean (p = .111) or 1 standard deviation above the mean of *SS* (p = .213).[17]

Hypothesis 4b predicts that with increasing *SS* sensitivity, seniors' detection of conclusion errors in the *Preparer Focus (documentation error)* condition will first increase, then decrease with increasing *SS* sensitivity. Results do not support the hypothesis. Neither the linear nor quadratic functions of SS are statistically significant (p > .104; one-tailed tests), see Table 2, Panel D. Tests of simple slopes reported in Table 3, Panel D are also not statistically significant (p = .101, .179, .216 at 1 standard deviation below, at, and above the mean of *SS*, respectively; one-tailed tests).

Auditor Rank and Errors Detected

In the analyses that follow, we test for differences in errors detected between managers and seniors. Specifically, with respect to documentation error detection when the preparer is expected to focus on conclusion errors, our theory suggests that managers but not seniors will experience mismatches in review templates at higher levels of *SS*. This implies that seniors will outperform managers in this context. With respect to conclusion error detection when the preparer focuses on documentation error, the reverse is true. Here, our theory suggests that seniors but not managers will experience review template mismatches at higher levels of *SS*, and, therefore, managers are expected to outperform seniors.

At low levels of *SS* (i.e., where auditors are relatively insensitive to stylization attempts), managers (seniors) are expected to use their default conclusion error (document error) review templates. Thus, at low levels of *SS*, managers will outperform seniors in terms of conclusion error detection because conclusion error detection is the default review template of managers (but not of seniors). Conversely, at low levels of *SS*, seniors will outperform managers in terms of documentation error detection because documentation error detection is the default review template of seniors (but not of managers).

Table 4 shows the predicted direction for the rank effects (positive sign implies managers outperform seniors; negative sign implies seniors outperform managers), along with the results. In the *Preparer Focus (conclusion error)* condition, we find support for two out of three directional predictions relating to seniors outperforming managers in terms of documentation errors detected, but we find no support for our prediction that managers outperform seniors in terms of conclusion errors detected at low levels of *SS*. In the *Preparer Focus (documentation error)* condition, results also support two out of three directional predictions relating to managers outperforming seniors in terms of conclusion errors detected, as well as the prediction that seniors outperform managers in terms of documentation errors detected at low levels of *SS*.

Auditor Rank and Knowledge Test

This section reports results of our tests of two maintained assumptions that we make in our framework: (1) that both seniors and managers have similar declarative technical

---

[17] This pattern is consistent with a concave upward-sloping curve, as suggested by the positive coefficient on the linear function that approaches significance, and the statistically significant negative coefficient on the quadratic function (Aiken and West 1991).

TABLE 4

Test of Rank Differences at Different Levels of Stylization Sensitivity (*SS*), with Documentation Errors and Conclusion Errors Detected as Dependent Variables

Panel A: Preparer Focus (Conclusion Error)

	Documentation Errors Detected as Dependent Variable			Conclusion Errors Detected as Dependent Variable		
	Predicted Sign	t	p-value	Predicted Sign	t	p-value
Rank effect at 1 standard deviation below mean of *SS*[a]	–	1.708	.048[b]	+	1.204	.118*
Rank effect at mean of *SS*	–	−1.643	.055*	none	.298	.767
Rank effect at 1 standard deviation above mean of *SS*	–	−3.439	.001*	none	−.995	.326

Panel B: Preparer Focus (Documentation Error)

	Documentation Errors Detected as Dependent Variable			Conclusion Errors Detected as Dependent Variable		
	Predicted Sign	t	p-value	Predicted Sign	t	p-value
Rank effect at 1 standard deviation below mean of *SS*	–	−2.453	.010*	+	1.307	.100*
Rank effect at mean of *SS*	none	−1.284	.207	+	1.533	.067*
Rank effect at 1 standard deviation above mean of *SS*	none	1.522	.136	+	.281	.390*

* one-tailed test.

[a] Rank effect refers to the extent to which managers outperform seniors. Thus, a statistically significant t-statistic with a positive (negative) sign means that managers outperform (underperform) seniors.

[b] Directionally opposite to our prediction that seniors will outperform managers.

knowledge relating to the inventory and accounts payable cycle (because these are aspects that both seniors and managers have frequent exposure to; Abdolmohammadi 1999); and (2) that managers (seniors) will have stronger (weaker) conclusion error review templates relative to documentation error review templates.

To verify the first assumption, we asked participants during the debriefing section to complete several technical knowledge questions pertaining to the inventory cycle and the associated accounts payable cycle. Results are consistent with our assumption of rank effects in terms of declarative technical knowledge (average standardized knowledge scores for seniors and managers = −0.043 and 0.048, respectively; p = .384).

To verify the second assumption, we presented participants with a stack of 20 randomly sorted cards describing review notes, half of which related to documentation errors, and half of which related to conclusion errors. We instructed participants to freely sort the cards into two piles on the basis that they found most appropriate (Chi et al. 1982; Nelson et al. 1995).

For each of the two piles of cards as sorted by each participant, we compute the proportion of cards that belongs to the dominant category. For example, for a participant who placed 13 cards (comprising 9 documentation and 4 conclusion errors) into one pile and 7 cards (comprising 6 conclusion and 1 documentation errors) in the other pile, the proportion for the documentation (conclusion) errors pile is computed as 9 divided by 13 (6 divided by 7). These proportions, therefore, measure the strength of participants' categorization of documentation and conclusion errors. In general, we find the sort proportions to be relatively high for both managers and seniors. For the sort proportion related to documentation errors, we find no statistically significant differences between managers and seniors (means = .911 vs. .884, respectively; p = .208; one-tailed test); however, for the sort proportion related to conclusion errors, managers outperform seniors (means = .941 vs. .884, respectively; p = .038; one-tailed test). To assess whether managers (seniors) have a stronger (weaker) conclusion error knowledge structure relative to documentation error knowledge structure, we also conduct a mixed ANOVA with the conclusion-error and documentation-error sort proportions as a within-subject factor, and rank as a between-subject factor. We find a marginally significant proportion by rank interaction (p = .067; one-tailed test). Among seniors, there is no significant difference in the relative strength of their conclusion-error and documentation-error sort proportions (mean = .884 for both proportions; p = .479, one-tailed test); among managers, their conclusion-error sort proportion is significantly higher than their documentation-error sort proportion (means = .941 vs. .911; p = .013, one-tailed test). Overall, these results are partially consistent with our second assumption. There is evidence that managers have stronger review templates pertaining to conclusion errors relative to documentation errors, and that their conclusion error review templates are also stronger than those of seniors.

V. CONCLUSION

In this paper, we present results that demonstrate that reviewers' responses to a preparer's expected stylization attempt (in terms of differential attention to documentation and conclusion errors) depend on three factors: nature of the stylization attempt, auditor rank, and stylization sensitivity. We show that when the preparer is anticipated to neglect documentation errors, seniors counteract this effect by detecting more documentation errors, and that documentation errors detected is an increasing function of their sensitivity to the preparer's stylization. In contrast, managers' detection of documentation errors first increases and then falls with increasing stylization sensitivity. Under the situation where the preparer is anticipated to neglect conclusion errors, we find evidence that managers serve as a control by detecting more conclusion errors with increasing stylization sensitivity. On the other hand, seniors' detection of conclusion errors is invariant with stylization sensitivity. The results have implications on the role of the review process, as they suggest that managers' (seniors') reviews may not serve as reliable controls for the detection of documentation (conclusion) errors under these settings.[18] Specifically, when a reviewer pays increasing attention toward error types potentially neglected by the preparer, and these error types are not consistent with the reviewer's template, performance in detecting that error may either remain unchanged (e.g., conclusion errors in the case of seniors) or even decrease (e.g., documentation errors in the case of managers).[19] Thus, there are limits to whether reviewers can successfully respond to preparers' expected stylization attempts.

---

[18] For example, Table 4 shows that in the ***Preparer Focus (conclusion error)*** condition, seniors outperform managers in documentation errors detection at the mean, and 1 standard deviation above the mean, of the ***Stylization Sensitivity*** variable; however, in the ***Preparer Focus (documentation error)*** condition, managers outperform seniors in conclusion errors detection at 1 standard deviation below, and at the mean, of the ***Stylization Sensitivity*** variable.

From a theoretical perspective, our study contributes to the audit expertise and review process literature by demonstrating that conclusions about rank-related differences in the review process need to take into account complex interactions between attributes of the reviewer and that of the preparer. Prior research shows that, in the review process, managers detect relatively more conclusion errors and fewer documentation errors than do seniors (Ramsay 1994). Our findings suggest that these rank differences are conditional on the particular combination of the preparer's expected stylization method and reviewer's stylization sensitivity. In addition, findings from our card-sorting task also extend the literature on reviewers' knowledge structures (e.g., Ramsay 1994) by providing more direct evidence on rank differences relating to the nature of auditors' review templates.

Prior research suggests that for reviewers to be effective, they must develop coping behaviors when faced with stylization (Rich et al. 1997). These coping behaviors are likely to be impacted by the reviewers' stylization sensitivity. In our study, we find differences in auditors' stylization sensitivity, and provide evidence that this sensitivity affects reviewers' performance depending on the nature of errors potentially neglected by preparers, and the match between these errors and the reviewers' underlying review templates. Furthermore, just as preparers are aware of the different preferences or focuses of different reviewers (Gibbins and Trotman 2002), it is expected that the preparers will, over time, learn the degree to which reviewers are sensitive to different types of stylization. Preparer's knowledge of reviewers' stylization sensitivity has implications on how they prepare audit workpapers. It may also have implications on how accountability pressures influence preparer behavior (Kennedy 1993; Tan 1995), which is a fruitful avenue for future research.

Our results showing the impact of a mismatch between auditor knowledge structures and error types on reviewer performance have wider implications. For example, the success of attempts by client personnel to act strategically (Rich et al. 1997) may be more or less successful, dependent on the sensitivity of auditors to such attempts, the nature of the error and account types, and the auditor's knowledge structure for these error and account types.

Several limitations of the paper and related research opportunities should be noted. First, we investigate a specific setting where the preparer is expected to stylize in terms of differential focus on either conclusion or documentation errors. While we believe that this form of stylization is important because of its direct impact on workpaper errors, there are clearly other ways that preparers can stylize workpapers (e.g., see Ricchiute 1999; Rich 2002). Second, the preparer's stylization approach is explicitly made known to the auditors participating in our experiment, and it is likely that in practice, auditors are less certain of the nature and extent of the stylization. Nonetheless, results from our debriefing questions indicate that auditors have observed different kinds of stylization in 30 percent to 40 percent of the audits they perform, which suggests that they do have some awareness of stylization attempts by preparers. Third, while our design choice to have audit seniors take over the role of the hypothetical ill manager likely leads to a bias against finding support for our hypotheses (see footnote 4), seniors' responses may be different if they had adopted the role of a senior. Fourth, the stylization sensitivity variable may be noisy because it is measured after the main experiment. We made this design trade-off because the alternative option of measuring the variable before the main experiment would likely lead to demand effects and contamination of the dependent variables. Finally, some CPA firms are also said to be moving toward an online review system. We believe that an online realtime review is likely to be more cognitively demanding on the reviewer, and may exacerbate (reduce) the negative (positive) effects of stylization sensitivity documented in the study. Mitigating this concern is that an online real-time review might also make it more cognitively difficult for the preparer to engage in stylization attempts.

---

[19] Bamber and Ramsay's (1997) basic finding that specialized reviews lead to poorer performance is consistent with the notion that dysfunctional performance arises when auditors focus exclusively on certain error types (and, therefore, deviating from their default review templates). They also find that managers (seniors) instructed to focus on conclusion (documentation) errors perform worse than those performing all-encompassing reviews. This finding is seemingly at odds with our results, which show that in the *Preparer Focus-documentation error* (*Preparer Focus-conclusion error*) condition, managers (seniors) with increasing stylization sensitivity and who focus more on detecting conclusion (documentation) errors actually perform better. Note that in Bamber and Ramsay's (1997) experiment, managers (seniors) do not receive any information on the heightened risk that the detection of conclusion (documentation) errors may be compromised. It is conceivable that awareness of this risk may alter their results. Another inference that may be drawn, in the context of our study, is that Bamber and Ramsay's (1997) results imply that, at the limit (when auditors focus exclusively on certain error types), dysfunctional performance can occur even though these error types are consistent with their default review templates. Of course, the contexts used in these prior studies and our study are different, and findings are not directly comparable. A fruitful avenue for future research would be to incorporate research designs that directly test these assertions.

REFERENCES

Abdolmohammadi, M. J. 1999. A comprehensive taxonomy of audit task structure, professional rank and decision aids for behavioral research. *Behavioral Research in Accounting* 11: 51–92.

Aiken, L. S., and S. G. West. 1991. *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage.

Bamber, E. M., and R. J. Ramsay. 1997. An investigation of the effects of specialization in audit workpaper review. *Contemporary Accounting Research* 14 (Fall): 501–513.

Biederman, I., A. L. Glass, and E. Stacey. 1973. Searching for objects in real world scenes. *Journal of Experimental Psychology* 97: 22–27.

Bower, G. H., J. B. Black, and T. J. Turner. 1979. Scripts in memory for text. *Cognitive Psychology* 11: 177–220.

Bruner, J. S. 1957. On perceptual readiness. *Psychological Review* 64: 123–152.

Chi, M. T. H., R. Glaser, and E. Rees. 1982. Expertise in problem solving. In *Advanced in Psychology of Human Intelligence*, edited by R. J. Sternberg, 7–75. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gibbins, M., and K. T. Trotman. 2002. Audit review: Managers' interpersonal expectations and conduct of the review. *Contemporary Accounting Research* 19 (3): 411–444.

Kahneman, D. 1973. *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall.

Kennedy, J. 1993. Debiasing audit judgement with accountability: A framework and experimental results. *Journal of Accounting Research* 31 (2): 231–245.

Libby, R., and K. T. Trotman. 1993. The review process as a control for differential recall of evidence in auditor judgment. *Accounting, Organizations and Society* 18 (6): 559–574.

McDaniel, M. A., and E. M. Kearney. 1984. Optimal learning strategies and their spontaneous use: The importance of task-appropriate processing. *Memory and Cognition* 12 (4): 361–373.

Nelson, M. W., R. Libby, and S. E. Bonner. 1995. Knowledge structure and the estimation of conditional probabilities in audit planning. *The Accounting Review* (January): 27–47.

Neter, J., W. Wasserman, and M. H. Kutner. 1996. *Applied Linear Statistical Models*. Chicago, IL: Irwin.

Ramsay, R. 1994. Senior/manager differences in audit workpaper review performances. *Journal of Accounting Research* 32 (Spring): 127–135.

Rehder, B. 2001. Interference between cognitive skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27 (2): 451–469.

Ricchiute, D. N. 1999. The effect of audit seniors' decisions on working paper documentation and on partners' decisions. *Accounting, Organizations and Society* 24: 155–171.

Rich, J. R. 2002. An experimental investigation of the review process from a persuasive perspective. Working paper, University of Connecticut.

Rich, J. S., I. Solomon, and K. T. Trotman. 1997. The audit review process: A characterization from the persuasion perspective. *Accounting, Organizations and Society* (May): 481–505.

Shiffrin, R. M., and W. Schneider. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84: 127–190.

Singley, M. K., and J. R. Anderson. 1989. *The Transfer of Cognitive Skill.* Cambridge, MA: Harvard University Press.

Tan, H-T. 1995. Effects of expectations, prior involvement and review awareness on memory for audit evidence and judgment. *Journal of Accounting Research* 33 (1): 113–135.

———, and J. Yip-Ow. 2001. Are reviewers' judgments influenced by memo structure and conclusions documented in audit workpapers? *Contemporary Accounting Research* 18 (4): 663–678.

Tanner, W. P., Jr., and J. A. Swets. 1954. A decision-making theory of visual detection. *Psychological Review* 61: 401–409.

Yip-Ow, J. and H-T. Tan. 2000. Effects of the preparer's justification on the reviewer's hypothesis generation and judgment in analytical procedures. *Accounting, Organizations and Society* 24: 203–215.

Editorial Data

The following table contains information about turnaround time for manuscripts (including revisions) on which editorial decisions were made in the 12-month period ended November 30, 2002. Turnaround time is the number of days between the date that the manuscript was received *and the date of the editor's letter to the author(s)*:

	Number of Manuscripts	Cumulative Number	Cumulative Percent
0 ≤ Days ≤ 30	67	67	18.41
31 ≤ Days ≤ 60	127	194	52.86
61 ≤ Days ≤ 90	120	314	86.26
91 ≤ Days ≤ 120	40	354	97.25
121 ≤ Days	10	364	100.00

The mean review time was 57.6 days; the median review time was 58 days.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal....No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of *The Chicago Manual of Style* (14th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's International Dictionary*.

FORMAT

1. All manuscripts should be typed in 12-point font on one side of 8½ × 11" good quality paper and be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form*. See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings.

Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix a, b, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says...."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by *The Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Date of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2nd edition. Homewood, IL: Dow Jones-Irwin.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, U.K.: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–636.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, The University of Texas at Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan or the like) should be submitted. Information that might identify the author(s) must be deleted from the instrument.
3. MS should be submitted via email as Microsoft® Word or PDF files to the *TAR* office at taruw@u.washington.edu. A submission letter, a separate cover page, and the MS containing no author identification should be emailed as separate files. The submission fee should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If unable to pay online, a check in U.S. funds, made payable to the American Accounting Association, may be sent to Professor Terry Shevlin, Senior Editor *TAR*, School of Business, Box 353200, University of Washington, Seattle, WA 98195-3200. The submission fee in U.S. funds is $75.00 for members and $100.00 for nonmembers of the AAA. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting. "An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

APRIL 2003 PLACEMENT ADS

The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The **Review** *is published on the first day of January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding.*

201 UNIVERSITY OF NOTRE DAME, Department of Accountancy, invites applications for position openings at the Assistant and Associate Professor levels. All research areas will be considered. Applicants for Assistant Professor positions should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong commitment to scholarly research and excellence in teaching. Salary is very competitive and research support is excellent. Please send *curriculum vitae* and an example of scholarly work to Professor Tom Schaefer, Chair, Department of Accountancy, Room 102D, Mendoza College of Business Administration, University of Notre Dame, Notre Dame, IN 46556-5646. The University of Notre Dame is an Equal Opportunity, Affirmative Action Employer.

202 ROLLINS COLLEGE, Crummer Graduate School of Business, invites applications for a tenure-track position at either the Associate or Full Professor level beginning in the Fall Semester 2003. The candidate, who will teach Financial and Managerial Accounting for M.B.A. students, will also teach elective courses. The Crummer School hires only experienced faculty with distinguished records of achievement. The candidate must demonstrate continued excellence in the classroom and publishing as prerequisites for tenure. We utilize the case method and make extensive use of computers in the classroom. *Forbes Magazine* ranks the Crummer School, located in the Orlando, Florida metropolitan area, as one of the nation's top 15 regional graduate business schools. The Crummer School, which is accredited by the AACSB, enrolls 400 students in four complementary M.B.A. program formats. Please send applications to Charles H. Brandon, Ph.D., Crummer Graduate School of Business, Rollins College, 2000 Holt Avenue, Winter Park, FL 32789. Rollins College is an Equal Opportunity Employer.

203 DUKE UNIVERSITY, Fuqua School of Business, is seeking to hire accounting faculty at both the junior and the senior levels. Hiring decisions will be based on the candidate's ability to contribute to the research and teaching missions of the Fuqua School of Business. Doctoral students interested in being considered for an Accounting faculty position at Duke should send a current resume, a paper or thesis proposal, and the names of references to Professor Jennifer Francis, Fuqua School of Business, Duke University, Durham, NC 27708. Materials should be sent by the middle of January 2003. Candidates for mid-career and senior appointments in Accounting should have significant research and teaching records. Persons who are interested in being considered for a mid-career or senior appointment in Accounting should send their materials to the above address. For more information, see our recruiting web site: http://faculty.fuqua.duke.edu/recruiting. Duke University is an Equal Opportunity, Affirmative Action Employer.

204 MONMOUTH UNIVERSITY, School of Business Administration, Department of Accounting and Business Law, invites applications for a tenure-track position at the Assistant/Associate Professor rank beginning Fall 2003. Applicants are required to have a Ph.D. at the time of appointment. (ABD will be considered.) All areas of teaching and research will be considered. Certification (CPA or CMA), prior teaching experience, and relevant practical experience are helpful. Interested parties should send a letter of application, *curriculum vitae,* and a list of three references to Dr. Nancy Uddin, School of Business Administration, 400 Cedar Avenue, West Long Branch, New Jersey 07764; Email: nuddin@monmouth.edu. Interested parties will also be interviewed at the 2003 Annual Meeting of the American Accounting Association. Contact Dr. Nancy Uddin. Applications will be considered until the position is filled. Visit our web site at http://www.monmouth.edu. Monmouth University is an Equal Opportunity, Affirmative Action Employer.

205 GEORGIA STATE UNIVERSITY, J. Mack Robinson College of Business, seeks candidates for one or more tenure- and nontenure-track faculty positions in the School of Accountancy to be effective Fall Semester 2003. Late-stage doctoral students, new Ph.D. holders, or experienced faculty members are encouraged to apply for a tenure-track position at the rank of Assistant Professor, with an interest in Information Systems. Candidates will be expected to conduct research leading to publication in the top refereed journals in Accounting and will be expected to perform well in the classroom. Excellent communication skills are required. Opportunities for conducting research in the School are outstanding. Salary is competitive and consistent with qualifications. Preference given to applications received by January 31, 2003. Positions are contingent on final budget approval by the University. Submit letter of application and resume, or nominations to Dr. A. Faye Borthick, Chair, Recruiting Committee, School of Accountancy, J. Mack Robinson College of Business, Georgia State University, PO Box 4050, 35 Broad Street, Suite 545, Atlanta, GA 30302-4050. Georgia State University is a unit of the University System of Georgia and is an Affirmative Action/Equal Opportunity Education Institution.

206 SHIPPENSBURG UNIVERSITY is seeking candidates for a tenure-track position in Accounting as Assistant Professor effective Fall 2003. Associate level may be considered, depending on qualifications. Ph.D. or D.B.A. in Accounting is required. ABD candidates may be considered. Terminal degree from an accredited institution is required for tenure. Position requires effective teaching, publishable research, service, and strong communication skills. Position may require teaching M.B.A. classes using videoconferencing technology. Qualified candidates should send letter of interest, *curriculum vitae,* undergraduate and graduate transcripts, and names, addresses, and telephone numbers of three references to Dr. Scott Cairns, John L. Grove College of Business, Shippensburg University, 1871 Old Main Drive, Shippensburg, PA 17257-2299. The College is accredited by AACSB. Shippensburg University is an Equal Opportunity Employer.

207 THE UNIVERSITY OF LOUISIANA AT MONROE, Department of Accounting, invites applications for two tenure-track positions, one in Auditing and one in Financial Accounting, for Fall 2003. The Accounting Program and the College of Business Administration are both accredited by the AACSB. Candidates must have an earned doctorate (or be ABD nearing completion); professional certification and experience are desirable. Applicants must have a commitment to excellence in teaching and scholarship. Salary and rank will be competitive and based on qualifications. Applicants should include a letter describing their ability to meet stated qualifications for the position, a *curriculum vitae,* photocopies of undergraduate and graduate transcripts, and the names of three references. Applications will be considered until the positions are filled and should be directed to Dr. James T. Hood, Head, Department of Accounting, The University of Louisiana at Monroe, Monroe, Louisiana 71209. For more information, visit the web site: http://www.ulm.edu. The University is an Equal Opportunity/Affirmative Action Employer located in a metropolitan area of 150,000 persons.

208 OAKLAND UNIVERSITY invites applications for a tenure-track, Assistant Professor of Accounting position. Applicants should have a Ph.D. degree in Accounting by September 2003, exhibit a willingness and ability to teach *Managerial Accounting* and *Accounting Information Systems* at the undergraduate and graduate levels, be committed to conduct scholarly research leading to publication in quality academic journals, and show a willingness to provide service to the university, school, and profession. Oakland University is a state-supported institution located 25 miles north of downtown Detroit in Rochester, Michigan. The School of Business Administration is AACSB-accredited. The normal teaching load is five classes per academic year. Class sizes are moderate. Research activity is supported and expected. For more information, see http://www.oakland.edu. If interested, please send an application package that includes your resume, evidence of your research potential, and evidence of your teaching capability to Edward J. Farragher, Chair, Department of Accounting and Finance, School of Business Administration, Oakland University, Rochester, MI 48309-4493; Email: farraghe@oakland.edu. Applications will be reviewed as they are received until the positions are filled. Oakland University is an Affirmative Action/Equal Opportunity Employer and encourages applications from women and minorities.

209 THE UNIVERSITY OF TEXAS AT DALLAS (UTD) invites applications for three positions in Accounting and Information Management effective September 2003. Appointment at Assistant, Associate, or Full Professor level will be considered. Appointees must have (or expect) a Ph.D. or equivalent and demonstrate the ability to conduct a nationally recognized research program. Graduate education in a university emphasizing research training and ability and interest in interacting with faculty colleagues on research projects in related disciplines are desirable. Applicants should send a complete *curriculum vitae* (educational history; teaching, research and work experience; list of publications, current research papers, and presentations), and names and addresses of three references to Academic Search #4100, The University of Texas at Dallas, PO Box 830688, M/S AD 23, Richardson, TX 75083-0688. For more information, visit http://www.utdallas.edu/dept/aim. UTD is an AA/EOE.

210 RYERSON UNIVERSITY, Toronto, Canada, invites applications for two tenure-track positions in the Accounting area in the School of Business Management, effective August 1, 2003. The successful candidates are expected to hold a Ph.D. or be in a Ph.D. program with a professional qualification. Preferably, the candidates will have a specialty in Assurance Services, Information System, Financial Accounting, or Taxation. Each appointment will be as an Assistant Professor or commensurate with the candidate's current status in another university. Salary and benefits are competitive. Located in the heart of Toronto, Ryerson is the ideal university to educate business leaders of tomorrow and to conduct strategic research in business disciplines. As a growing university, Ryerson offers great opportunities for academic career achievements. Applicants should submit by email or by mail their *curriculum vitae* (and teaching evaluations, if any) to Dr. Peter Luk, Director, School of Business Management, Ryerson University, 350 Victoria St., Toronto, Ontario, M5B 2K3 Canada; Email: peterluk@ryerson.ca.

211 SIENA COLLEGE, Accounting Department, invites applications from experienced teacher/scholars for two tenure-track positions—beginning Fall Semester 2003. Applications will be accepted until the positions are filled. Rank and salary are dependent upon experience and credentials. The positions require a Ph.D. or a D.B.A. with a concentration in Accounting. Candidate(s) for a senior-level appointment should be able to demonstrate a track record of teaching excellence and a sustained record of professional development activities. Relevant business experience and professional licensing are preferred. Experience and research interests in Auditing, Managerial Accounting, or Accounting Systems are highly desirable. Siena College is an independent liberal arts college in the Franciscan tradition, located in New York's Capital District, in close proximity to other major higher education and cultural institutions. The college enrolls 2,800 full-time students, 1,100 in the School of Business. The School of Business is currently in candidacy for AACSB accreditation. Please forward applications to Accounting Search Committee, C/O School of Business, 515 Loudon Road, Loudonville, NY 12211. Siena College is an Equal Opportunity Employer and encourages applications from all qualified candidates.

212 UNIVERSITY OF HOUSTON, C.T. Bauer College of Business, Department of Accountancy and Taxation, will interview candidates for tenure-track appointments at any rank for the 2003–2004 academic year. A doctoral degree (or near completion) and a strong commitment to both research and teaching are required. We are seeking applicants in the areas of Accounting Information Systems, Auditing, Financial Accounting, Managerial Accounting, and Taxation. The Bauer College of Business and the department are AACSB-accredited. Salaries are comparative. Please send a copy of your *curriculum vitae* and a research paper to Recruiting Committee, c/o Dr. George O. Gamble, University of Houston, C.T. Bauer College of Business, 334 Melcher Hall, Houston, TX 77204-6021, or by email to ggamble@uh.edu. The University of Houston in an Equal Opportunity/Affirmative Action Employer. Women, minorities, veterans and individuals with disabilities are encouraged to apply.

213 COASTAL CAROLINA UNIVERSITY, E. Craig Wall, Sr. College of Business, is seeking a qualified applicant at the Assistant Professor level to teach Financial/Managerial Accounting Principles and a specialization in one or more of the following areas: AIS, Auditing, Cost, or Governmental. The successful candidate should possess a Ph.D./D.B.A. in Accounting. Candidates who are AACSB academically qualified or ABDs near completion are encouraged to apply. Some business experience and/or professional certification is desirable. The search will continue until the position is filled. Coastal Carolina University, a public, predominantly undergraduate liberal arts institution located nine miles from Myrtle Beach, South Carolina, enrolls more than 5,800 students from 48 states and 50 countries. The University offers baccalaureate degree programs in 36 major fields of study and master's degrees in Marine Science and Education. Please send application letter and resume to Professor James F. Eason, Associate Dean; Wall College of Business; Coastal Carolina University; PO Box 261954; Conway, SC 29528-6054; Email: Jim@Coastal.edu. Coastal Carolina University is an EO/AA Employer.

214 UNIVERSITY OF ALASKA FAIRBANKS, School of Management, invites applications for the position of Professor/Associate of Accounting. Both the School of Management and Accounting Program are accredited by AACSB. The successful candidate is expected to offer high-quality instruction in area of expertise and to make a major contribution to Department of Accounting/Information Systems, providing leadership in teaching, research, and service for the Accounting Program. Responsibilities will include playing a major role in the further development of the program and acting as a principal catalyst to strengthen relationships between the Department and professional community. Candidates must hold a Ph.D. in Accounting or related field and provide evidence of teaching, research, and communications skills; CPA, CMA, or CIA certification preferred. Salary is competitive and commensurate with qualifications. Applications will be reviewed on March 1, 2003, and position will be open until filled. Send a signed UAF application form available from http://www.uaf.edu/uafhr; cover letter; *curriculum vitae;* three letters of reference; evidence of teaching effectiveness; and examples of scholarly activity to Dean James Collins, c/o UAF Human Resources, PO Box 757860, Fairbanks, AK 99775-7860.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 611–639

Do Nonaudit Services Compromise Auditor Independence? Further Evidence

Hollis Ashbaugh
Ryan LaFond
Brian W. Mayhew
University of Wisconsin–Madison

ABSTRACT: This paper challenges the findings of Frankel et al. (2002) (FJN). The results of our discretionary accruals tests differ from FJN's when we adjust discretionary current accruals for firm performance. In our earnings benchmark tests, in contrast to FJN we find no statistically significant association between firms meeting analyst forecasts and auditor fees. Our market reaction tests also provide different results than those reported by FJN. Overall, our study indicates that FJN's results are sensitive to research design choices, and we find no systematic evidence supporting their claim that auditors violate their independence as a result of clients purchasing relatively more nonaudit services.

Keywords: *independence; audit fees; discretionary accruals; biased financial reporting.*

Data Availability: *Data collected from public sources. Authors will make sample firms available upon request.*

I. INTRODUCTION

Regulators, financial statement users, and researchers are concerned that auditors compromise their independence by allowing high fee clients more financial statement discretion relative to low fee clients. Frankel et al. (2002) (hereafter FJN) use the association between audit firm fees and two measures of biased financial reporting—firms' discretionary accruals and the likelihood of firms meeting earnings benchmarks—to draw inferences on auditor independence. In addition, they conduct an event study to investigate whether the market reacts to the disclosure of auditor fees. FJN suggest that their results provide evidence that auditor independence is compromised when clients pay high nonaudit fees relative to total fees. We conduct the same three sets of empirical tests as FJN to investigate the sensitivity of FJN's results to research design choices.

The motivation for FJN's study and our own is based on the U.S. Securities and Exchange Commission's (SEC) concern about the growth of nonaudit fees relative to audit

---

We appreciate the comments received at the 2002 American Accounting Association Annual Meeting, 2002 University of Illinois Audit Symposium, University of Missouri, University of Wisconsin–Madison, and from Jere Francis, Karla Johnstone, Bill Kinney, Joel Pike, and Terry Warfield. We acknowledge the financial support in collecting the data for this study provided by the Department of Accounting at the University of Wisconsin–Madison and its many supporters.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted April 2002
Accepted January 2003

fees during the 1990s (e.g., see Levitt 2000). The SEC's concern that the growth in the provision of nonaudit services compromises audit firm independence is based on the premise that the provision of nonaudit services increases the fees paid to the audit firm, thereby increasing the economic dependence of the audit firm on the client. Prior research agrees that it is the strength of the economic bond between the audit firm and its client that reduces auditor independence (DeAngelo 1981; Beck et al. 1988; Magee and Tseng 1990). Yet, there is dissention on how to measure the economic bond. Much of the contemporaneous research investigating whether the provision of nonaudit services decreases auditor independence uses the ratio of nonaudit fees to total fees (hereafter referred to as the fee ratio) as the measure of the economic bond.[1] However, the fee ratio does not capture the economic importance of the client to the audit firm when the total client fees are immaterial to the audit firm.

Consider the following: two firms in our sample have a fee ratio of 73 percent; one firm reports total fees of $71,000 and the other firm reports total fees of $5.7 million. Based on their fee ratios, both firms are considered threats to independence, whereas only the latter is economically significant to the auditor. Thus, we posit that the total fee, i.e., the sum of audit and nonaudit fees, rather than the fee ratio is the more appropriate measure of the economic dependence of the auditor on a client, although, we include the fee ratio in our empirical analyses to provide a complete assessment of FJN's results.

Our first set of tests examine whether FJN's discretionary accruals results are sensitive to the measurement of discretionary accruals. Recent research suggests that the discretionary accrual measure used by FJN systematically rejects the null hypothesis of no-biased financial reporting at much higher-than-stated rates because the measure does not account for the impact of performance on accruals (Kothari et al. 2002). We calculate two measures of discretionary current accruals that control for the impact of firm performance in the estimation of discretionary accruals. The first method measures performance-adjusted discretionary current accruals as the difference between a firm's estimated discretionary current accruals and the median discretionary current accruals from an industry- and performance-matched portfolio. The second approach follows Kothari et al. (2002) and includes a control for firm performance in the regression model used to estimate nondiscretionary current accruals.

Consistent with FJN, we document a positive association between the absolute value of firms' discretionary current accruals and fee ratio, regardless of which performance-adjusted accrual measure is used. We also find, like FJN, no evidence that firms' total fees are associated with firms' discretionary current accruals. However, in contrast to FJN, when we partition sample firms based on whether they report income-increasing or income-decreasing discretionary accruals, and we employ our performance-adjusted measures of current discretionary accruals, we find no association between the fee ratio and income-increasing discretionary accruals. Further analysis suggests the measurement error caused by not controlling for firm performance in the estimate of income-increasing discretionary accruals is associated with the fee ratio. Thus we provide an explanation for why FJN find an association between the fee ratio and income-increasing discretionary accruals, whereas we do not.

---

[1] The use of a fee ratio may be, in part, driven by prior research that relied on Accounting Series Release (ASR) No. 250 fee disclosures (e.g., see Beck et al. 1988; Read and Tomczyk 1992). In 1978, the SEC issued ASR No. 250, which required U.S. publicly traded companies to disclose the types of nonaudit services purchased from incumbent auditors expressed as a percentage of their total audit fees.

Our results indicate that the association between the fee ratio and the absolute value of discretionary accruals is driven by income-decreasing discretionary accruals. While income-decreasing accruals can be interpreted as a form of biased financial reporting, income-decreasing accruals also reflect a conservative application of generally accepted accounting principles (GAAP). Typically, regulators and financial statement users are more concerned with the opportunistic application of GAAP than the conservative application of GAAP, with the opportunistic application of GAAP more likely signaling problems with auditor independence (Becker et al. 1998).[2]

FJN also use the relation between audit firm fees and two earnings benchmarks—small earnings increases and meeting analyst forecasts—to draw inferences on audit firm independence. Our second set of tests replicate FJN's earnings benchmark tests. Consistent with FJN, as well as with Francis and Ke (2001), we find no association between fee ratio and the likelihood of firms reporting small earnings increases. In addition, we, like FJN, document a negative association between total fees and the likelihood of firms reporting small earnings increases. However, unlike FJN, we find no statistically significant association between either the fee ratio or total fees, and firms meeting analyst forecasts.

In summary, based on the use of discretionary accruals and earnings benchmarks as proxies for biased financial reporting, we find little evidence supporting the claim that auditors violate their independence as the result of clients paying high fees or having high fee ratios. Our discretionary accrual analyses appear consistent with concurrent research examining audit firm fees and measures of earnings management. Chung and Kallapur (2003) find no association between their audit fee metrics and the absolute value of discretionary accruals measured with the modified Jones model.[3] They also explore four different subsets of firms where they expect firms to have greater incentives to manage earnings (i.e., firms with high market-to-book ratios, high leverage, low audit-firm industry specialization, and high insider influence on the board of directors) and find no evidence of an association between audit or nonaudit fees and earnings manipulation. The inferences drawn from our discretionary accrual and earnings benchmark tests are also consistent with those reported by DeFond et al. (2002) who show no association between firms' going concern opinions and magnitude of nonaudit services, and a positive association between the likelihood of issuing a going concern opinion and audit fees.

In addition to supporting the findings of concurrent audit research, the results of our discretionary accrual analysis extend the contemporaneous literature that addresses the estimation of firms' discretionary accruals. Our results indicate that controlling for performance is important when estimating managers' discretionary reporting practices. We also provide empirical evidence that measurement error in discretionary accruals can lead to erroneous inferences when the measurement error is correlated with the test variable.

The last set of tests that we conduct investigates the market's reaction to the disclosure of auditor fees in firms' proxy statements. FJN report that they find a significant negative association between share prices and the disclosure of higher than expected nonaudit fees, albeit they state the association is small in economic terms and insignificant when measured over long event windows. We have two primary concerns related to FJN's market reaction analysis. First, when we calculate abnormal returns on firms' proxy filing dates, we find that the one-day market-adjusted return is less than one half of 1 percent, and is not significantly different from zero. Our second concern is that firms are required to make a substantial number of disclosures in their proxy statements. Efforts to examine the relation between fee metrics and changes in market value without controlling for the other information contained in proxy statements can result in incorrect inferences being drawn on whether the market reacts to the disclosure of auditor fees. When we conduct market reaction tests examining the difference in abnormal returns on the filing of a firm's 2000 versus 1999 proxy statements, unlike FJN we find no evidence that the market, on average, reacts to the fee ratio. We posit that it is difficult to draw reliable conclusions about the market's reaction to auditor fee disclosures without controlling for the firm-specific disclosures contained in proxy statements.

---

[2] One exception is when firms, through the overly conservative application of GAAP, construct reserves that are used in future periods to inflate or smooth earnings (see e.g., Levitt 1998).

[3] Chung and Kallapur (2003) use the ratio of client audit fee to total audit firm revenues and client nonaudit fee to total audit firm revenue as measures of the threat to auditor independence. In addition, they use client total fees and nonaudit fees divided by an estimate of practice office revenues as proxies for threats to auditor independence.

The paper proceeds as follows. The next section discusses the alternative ways to combine fee data to capture threats to auditor independence. Section III describes our sample and identifies the determinants of auditor fees that we need to control for in our empirical tests. Section IV presents the research design and results. The last section summarizes our findings relative to those of FJN.

II. MEASURING THREATS TO AUDITOR INDEPENDENCE

DeAngelo (1981) models that, as the economic bond between the audit firm and client increases, the audit firm's dependence on the client increases.[4] Nonaudit fees further increase the client-auditor bond by increasing the portion of audit-firm wealth derived from a client (Simunic 1984; Beck et al. 1988). Nonaudit fees can also threaten independence when clients use them as contingent fees. Magee and Tseng (1990) note that while contingent fees are explicitly prohibited by audit standards, clients can create contingent fees by withholding profitable nonaudit services when the auditor does not allow the client to report its preferred financial condition. We argue that the sum of audit and nonaudit fees, i.e., total fees, best captures the explicit economic bond between the audit firm and client. We do not advocate the use of the fee ratio because the fee ratio does not necessarily capture the economic importance of the client to the audit firm. The fee ratio does, however, capture the relative monetary value of the audit versus nonaudit services provided by the audit firm to a client, which may have an impact on the perception of independence held by regulators, as well as the general public. As stated earlier, we include the fee ratio in our empirical analyses to be comparable to FJN and contemporaneous auditor independence research.

Our research design, like FJN, investigates whether the degree of client auditor bonding is associated with evidence that the financial statements reflect a biased view of the client's financial condition. We both use the magnitude of discretionary accruals and the likelihood of meeting earnings benchmarks as evidence of biased financial statements. Discretionary accruals and earnings benchmarks can be viewed as evidence of auditor independence violations based on three assumptions. The first assumption is that an independent auditor requires his/her client to file objective (i.e. unbiased) financial statements. Second, discretionary accruals provide a measure of the degree of bias inserted into the financial statements by management and allowed by the auditor. Third, just meeting or beating earnings benchmarks such as prior years' earnings or current analyst forecasts is not random, suggesting that managers bias financial statements to meet earnings targets (see Burgstahler and Dichev 1997). An objective auditor will not allow clients to bias earnings to achieve these benchmarks even as his economic dependence on the client increases.

---

[4] We view the audit firm as the U.S. firm. As part of our empirical analyses, we conduct two sensitivity tests to assess whether there are specific audit firm effects. First, we estimate audit firm-specific regressions. Second, we conduct categorical analysis incorporating binary variables for each audit firm. We find no specific audit firm effects.

III. AUDITOR FEES

Sample and Descriptive Statistics

We collected fee data for 4,959 firms from U.S. registrants' 2000 proxy statements that were available on EDGAR or Global Access during November and December 2001. We deleted 761 financial firms and 1,028 firms that lacked the necessary financial data on Compustat to estimate our fee models arriving at a final sample of 3,170 firms.

Panel A of Table 1 reports descriptive statistics for the four audit fee metrics that we use in our empirical tests. *AUDIT* and *NONAUDIT* are the audit and nonaudit fees reported in sample firms' proxy statements, respectively.[5] *TOTAL* is the sum of *AUDIT* and *NONAUDIT*. *FEERATIO* is defined as *NONAUDIT* divided by *TOTAL*. We also report fee metrics partitioned by Big 5 and non-Big 5 clients, where Big 5 is defined as the audit firms of Arthur Andersen, Deloitte & Touche, Ernst & Young, KPMG, or PricewaterhouseCoopers. We partition the sample by auditor choice because our fee data indicate that the magnitude of audit and nonaudit fees differ between Big 5 and non-Big 5. The mean (median) of *AUDIT* for Big 5 clients is \$562,420 (\$212,000) and the mean (median) of *NONAUDIT* is \$1,332,408 (\$240,100). In contrast, the mean (median) of *AUDIT* for non-Big 5 clients is \$149,371 (\$88,750) and the mean (median) of *NONAUDIT* is \$136,483 (\$33,360), resulting in *TOTAL* for non-Big 5 auditors being smaller than for Big 5 auditors. The descriptive statistics also indicate that the mean (median) *FEERATIO* is significantly smaller for non-Big 5 auditors (30.32 percent [27.87 percent]) than Big 5 auditors (49.60 percent [51.90 percent], $p < .001$).

To better understand the relation between client size and the demand for audit-related services, we report in Panel B of Table 1 the mean (median) values of *FEERATIO*, *TOTAL/ASSETS*, *BIG5*, *ASSETS*, and *MVE* across *TOTAL* deciles. *MVE* is equal to a firm's market value of equity defined as price per share at fiscal year end (Compustat data item 199) multiplied by the number of shares outstanding (Compustat data item 25) measured in millions of dollars. *TOTAL* is directly related to client size. *FEERATIO*, *ASSETS*, and *MVE* are increasing in *TOTAL*, and the proportion of firms contracting with a Big 5 auditor is increasing in *TOTAL*.

Panel C of Table 1 reports the distribution of our sample firms across the industry classifications reported in FJN, and shows that our sample's industry composition is closely aligned to the industry composition in the Compustat database. Our sample's descriptive statistics are similar, in all respects, to FJN's.

Fee Determinants

We investigate the determinants of auditor fees for two reasons. First, 2000 is the first year that auditor fee data has been widely available for U.S. publicly traded companies. We want to establish that the determinants of auditor fees identified by prior research using fee data from foreign countries, as well as U.S. data collected via surveys, are valid for our sample. Second, we want to determine whether variables that prior research documents are correlated with discretionary accruals are also correlated with auditor fees. Identification

---

[5] The SEC requires separate disclosure of information system design and implementation fees. We include these fees in *NONAUDIT*.

TABLE 1
Descriptive Statistics of Audit and Nonaudit Fees

Panel A: Alternative Fee Metrics

	Mean	**Median**	**STD**
Full sample (n = 3170)			
AUDIT	514,601	190,000	1,517,051
NONAUDIT	1,193,952	197,722	4,646,406
TOTAL	1,708,553	417,180	5,802,349
FEERATIO	47.37%	48.28%	0.24
Sample firms audited by the Big 5 (n = 2803)			
AUDIT	562,420	212,000	1,604,922
NONAUDIT	1,332,408	240,100	4,921,499
TOTAL	1,894,828	488,000	6,141,965
FEERATIO	49.60%	51.90%	0.24
Sample firms audited by Non-Big 5 (n = 367)			
AUDIT	149,371	88,750	237,193
NONAUDIT	136,483	33,360	479,088
TOTAL	285,854	135,967	639,214
FEERATIO	30.32%	27.87%	0.21

The difference between the means (medians) of Big 5 and non-Big 5 are significant at the .001 level.

Panel B: Means (Medians[in italics]) Ranked by TOTAL

Group	***TOTAL***	***FEERATIO***	***TOTAL/ASSETS***	***BIG5***	***ASSETS***	***MVE***
1	66,016	20.81%	0.58%	0.58	51	79
	70,000	*20.78%*	*0.40%*		*16*	*14*
2	123,571	27.76%	0.78%	0.77	66	77
	125,333	*26.97%*	*0.37%*		*34*	*26*
3	182,853	34.31%	0.48%	0.83	120	169
	185,000	*34.42%*	*0.29%*		*64*	*47*
4	254,881	39.70%	0.35%	0.88	228	369
	255,490	*39.82%*	*0.22%*		*118*	*86*
5	354,859	45.55%	0.35%	0.92	297	393
	354,000	*47.21%*	*0.21%*		*157*	*117*
6	498,551	52.26%	0.37%	0.92	456	612
	498,300	*54.23%*	*0.23%*		*216*	*190*
7	719,558	55.79%	0.34%	0.97	737	1031
	713,000	*56.81%*	*0.20%*		*352*	*329*
8	1,116,605	60.70%	0.28%	0.97	1153	1330
	1,100,000	*63.15%*	*0.17%*		*610*	*571*
9	2,116,894	65.20%	0.22%	0.98	2926	3422
	2,025,990	*67.73%*	*0.14%*		*1571*	*1365*
10	11,681,726	71.54%	0.27%	0.99	17351	20903
	6,344,008	*73.62%*	*0.11%*		*7538*	*5433*
Mean	1,708,553	47.37%	0.40%	0.88	2334	2834
Median	*417,180*	*48.38%*	*0.22%*		*209*	*166*

(*continued on next page*)

some economic event or management choice. The portfolio approach controls for relative firm performance across random samples.[10]

Our second measure of discretionary accruals, *REDCA*, follows Kothari et al. (2002) by including lagged *ROA* in the accruals regression to control for firm performance. Kothari et al. (2002) report that matching on lagged *ROA*, as opposed to contemporaneous *ROA*, eliminates any mechanical relation between current period's discretionary accrual estimate and the performance metric. The calculation of *REDCA* begins by estimating the following cross-sectional current accrual regression by each two-digit SIC code partition:

$$CA = \gamma_1(1/laglasset) + \gamma_2(\Delta\ \mathrm{Rev}) + \gamma_3 Lag1ROA + \varepsilon \tag{4}$$

The parameters from Equation (4) are used to calculate expected current accruals estimated with a performance control (*ECAPC*):[11]

$$ECAPC = \hat{\gamma}_1(1/Laglasset) + \hat{\gamma}_2(\Delta\ \mathrm{Rev} - \Delta\ AR) + \hat{\gamma}_3 Lag1ROA \tag{5}$$

REDCA is equal to *CA* minus *ECAPC*. In contrast to *PADCA*, which controls for relative firm performance within two-digit SIC classes, the *REDCA* estimate of discretionary current accruals controls for performance on a firm-specific basis.[12]

Table 3 reports descriptive statistics on the variables used in our empirical tests for the 3,069 sample firms having data to calculate *PADCA* and *REDCA*.[13] The mean (median) values of *PADCA* and *REDCA*, both of which are deflated by *ASSETS*, are −1.08 percent (0.08 percent) and −0.66 percent (0.53 percent), respectively. The means of *PADCA* and *REDCA* are significantly different at the .001 level. Panel B of Table 3 reports the correlations between the alternative measures of discretionary accruals. While highly correlated, the correlations between *DCA*, *PADCA*, and *REDCA* are significantly less than 1, indicating each variable is a distinct estimate of a firm's financial statement bias.

Discretionary Accrual Model

We use the following OLS regression model to test the stability of FJN discretionary accrual results:

---

[10] A caveat to the portfolio approach is that it may result in a non-random match if all firms in the portfolio manage their accruals and do so at identical levels. This would decrease the likelihood of finding an association between discretionary accruals and auditor fees. However, within each of the portfolios there is variation in firms' discretionary accruals and as a consequence *PADCA* captures a relative measure of manipulation. Our *REDCA* measure is not subject to this concern.

[11] The mean R^2s of Equation (2) and Equation (4) are 0.11 and 0.18, respectively, suggesting that the CA model that includes lagged ROA is a better model of firms' expected accruals.

[12] The advantage of including lagged ROA in the estimation of discretionary accruals is that it allows for the association between firm performance and accruals to differ across industries. An alternative to including lagged ROA in the estimation of discretionary accruals is to include it as an explanatory variable in the discretionary accrual regression. The disadvantage of this approach to controlling for firm performance is it assumes that lagged ROA has a constant association to discretionary accruals across industries.

[13] We require at least 20 firms in each two-digit SIC partition to estimate discretionary accruals. This results in a sample reduction of 101 firms.

TABLE 3
Descriptive Statistics for Alternative Measures of Discretionary Current Accruals

Panel A: Variables Used in Discretionary Accruals Analysis (n = 3069)

	Mean	Median	STD
PADCA	−1.08%	0.08%	0.23
REDCA	−0.66%	0.53%	0.25
L1ACCRUAL	−7.05%	−0.15%	0.36
AUDIT	0.51	0.19	1.52
NONAUDIT	1.21	0.20	4.69
TOTAL	1.72	0.41	5.84
FEERATIO	47.45%	48.39%	0.24
ASSETS	2,341	205	13,761
BIG5	0.88	1.00	0.32
MVE	2,831	169	16,662
MERGER	0.24	0.00	0.43
FINANCING	0.33	0.00	0.47
LEVERAGE	0.47	0.48	0.25
MB	3.49	1.77	8.00
LITIGATION	0.40	0.00	0.49
INST_HOLDING	32.98%	28.63%	0.28
LOSS	0.43	0.00	0.50
CFO	−0.04	0.05	0.43

Panel B: Pearson (Spearman) Correlations Above (Below) the Diagonal of Alternative Discretionary Accrual Measures[a]

	DCA	*PADCA*	*REDCA*
Discretionary Current Accruals (*DCA*)		0.92 *0.00*	0.93 *0.00*
PADCA[b]	0.88 *0.00*		0.91 *0.00*
REDCA[b]	0.86 *0.00*	0.89 *0.00*	

---

[a] p-values of the correlations are reported in italics.

[b] All variables used in calculating *PACDA* and *RECDA* are winsorized at the 1 and 99 percent, respectively.

PADCA = the discretionary current accruals measure controlling for performance using the portfolio match technique;

REDCA = the discretionary current accruals measure controlling for performance by including the prior years' ROA in the estimation of expected accruals;

L1ACCRUAL = last year's total current accruals equal to net income before extraordinary items (Compustat data item 123) plus depreciation and amortization (Compustat data item 125) minus operating cash flows (Compustat data item 308) scaled by beginning of year total assets.

AUDIT = the audit fee in millions of dollars;

NONAUDIT = the nonaudit fee in millions of dollars;

TOTAL = the sum of *AUDIT* and *NONAUDIT*;

FEERATIO = *NONAUDIT* divided by *TOTAL*;

ASSETS = a firm's total assets (Compustat data item 6) in millions of dollars;

BIG5 = 1 if the sample firm is audited by Arthur Andersen, Deloitte & Touche, Ernst & Young, KPMG, or PricewaterhouseCoopers (identified by Compustat data item 149), and 0 otherwise;

(*continued on next page*)

TABLE 3 (continued)

MVE = a firm's price per share at fiscal year end (Compustat data item 199) multiplied by the number of shares outstanding (Compustat data item 25) measured in millions of dollars;
MERGER = 1 if the sample firm has engaged in a merger or acquisition (identified by Compustat AFTNT1), and 0 otherwise;
FINANCING = 1 if merger is not equal to 1 and number of shares outstanding increased by at least 10 percent, or long-term debt increased by at least 20 percent, or the firm first appeared on the CRSP monthly returns database during the 2000 fiscal year, and 0 otherwise;
LEVERAGE = a firm's total assets less its book value (Compustat data item 60) divided by its total assets;
MB = a firm's market-to-book ratio defined as its market value of equity divided by book value;
LITIGATION = 1 if the firm operates in a high-litigation industry, and 0 otherwise (high-litigation industries are industries with SIC codes of 2833–2836, 3570–3577, 3600–3674, 5200–5961, and 7370–7370;
INST_HOLDING = the percentage of shares held by institutional owners at the beginning of the 2000 calendar year;
LOSS = 1 if the firm reports a net loss in fiscal 2000, and 0 otherwise; and
CFO = cash flow from operations, defined as Compustat data item 308 scaled by beginning of year total assets.

$$\begin{aligned} DCA\_PA = {} & \alpha + \beta_1 FEE + \beta_2 BIG5 + \beta_3 L1ACCRUAL + \beta_4 \ln MVE \\ & + \beta_5 MERGER + \beta_6 FINANCING + \beta_7 LEVERAGE + \beta_8 MB \\ & + \beta_9 LITIGATION + \beta_{10} INST\_HOLDING + \beta_{11} LOSS \\ & + \beta_{12} CFO + \varepsilon \end{aligned} \tag{6}$$

where:

DCA_PA = the discretionary current accrual measure adjusted for firm performance, which is set equal to *PADCA* or *REDCA*;
FEE = *FEERATIO*, or the natural log of *TOTAL*, *AUDIT*, or *NONAUDIT*;
L1ACCRUAL = last year's total current accruals equal to net income before extraordinary items (Compustat data item 123) plus depreciation and amortization (Compustat data item 125) minus operating cash flows (Compustat data item 308) scaled by beginning of year total assets;
LITIGATION = 1 if the firm operates in a high-litigation industry, and 0 otherwise. High-litigation industries are industries with SIC codes of 2833–2836, 3570–3577, 3600–3674, 5200–5961, and 7370–7370;
INST_HOLDING = the percentage of shares held by institutional owners at the beginning of 2000 calendar year;
LOSS = 1 if the firm reports a net loss in fiscal year 2000, and 0 otherwise;
CFO = cash flow from operations, defined as Compustat data item 308 scaled by beginning of year total assets; and

All other variables are as previously defined.

Equation (6) is different from the model that FJN use to test the association between auditor fees and discretionary accruals in several ways. First, FJN use total accruals and the absolute value of total accruals in their regression model to control for firm performance. We do not include total accruals or the absolute value of total accruals in our model because these variables are used to derive discretionary accruals, which results in an overstatement of the explanatory power of the model.[14] We do, however, add the prior year's current accruals (*L1ACCRUAL*) to the model to capture the reversal of accruals over time.

Second, FJN use one variable to capture firms' business combinations and financing activities. We use two separate variables, *MERGER* and *FINANCING*, because the two types of transactions may have different consequences for firms' accruals, and the results of our fee analyses indicate different relations between audit-related fees and these two economic events. Third, we drop audit tenure from the FJN model, as this variable does not add any explanatory power to our model.[15] We also exclude the absolute value of cash flows from operations from the model, as including this variable induces multicollinearity with the cash flow from operations variable. Finally, we transform the dependent variable, *PADCA* or *REDCA*, by taking the natural log of the absolute value to correct for violations of normality (Warfield et al. 1995).

Discretionary Accruals Results

Table 4 reports the results of our unsigned discretionary accruals tests. We find a positive relation between *FEERATIO* and the absolute value of *PADCA* and the absolute value of *REDCA*. We find no significant association between *TOTAL* and either unsigned estimate of discretionary accruals. The results on *AUDIT* and *NONAUDIT* are conditional on the discretionary accrual estimate used. The coefficient on *NONAUDIT* is positive and significant only in the *PADCA* model, whereas the coefficient on *AUDIT* is negative and significant only in the *REDCA* model. The coefficients on the control variables are, in general, significant and stable regardless of fee metric or discretionary accruals measure. The sign and significance of the coefficients on *FEERATIO* are consistent with FJN's results regardless of the discretionary accrual estimate used. In contrast to FJN, however, we find the significance of the coefficients on *NONAUDIT* and *AUDIT* to be sensitive to the estimate of discretionary accruals used.

To provide evidence on whether there is any differential relation between the fee metrics and our measures of discretionary accruals conditional on whether discretionary accruals are income-increasing or income-decreasing, we partition the sample into two groups based on the sign of firms' discretionary accruals. Panel A of Table 5 reports the regression results estimated using sample firms that report *PADCA* or *REDCA* greater than or equal to zero. We find no significant association between *FEERATIO* and positive *PADCA* or *REDCA* measures. Likewise, our results indicate no significant association between *TOTAL* and positive *PADCA* or *REDCA* measures. Furthermore, we find no evidence that *AUDIT* or *NONAUDIT* is significantly related to positive discretionary accruals. These findings are in contrast to FJN's results, where they report *FEERATIO*, *NONAUDIT*, and *TOTAL* to be positively related to positive discretionary accruals and *AUDIT* to be negatively related to positive discretionary accruals.

Panel B of Table 5 reports the regression results estimated using sample firms having *PADCA* or *REDCA* values less than zero. We find a negative relation between *FEERATIO*

---

[14] For example, when we estimate the discretionary accrual model that includes total accruals and the absolute value of total accruals and uses *FEERATIO* as the fee metric, the adjusted R^2 is 43 percent. Re-estimating the model excluding these two variables results in an R^2 of 16 percent.

[15] In contemporaneous research, Myers et al. (2003) examine auditor tenure and discretionary accruals. Myers et al. (2003) note that when conducting sign specific breakdowns of discretionary accruals OLS estimates are in general biased toward zero, due to the fact that this analysis constrains the distribution of the dependant variable. To address this potential problem Myers et al. (2003) use a maximum likelihood truncated regression approach when estimating their models for the signed analysis. When we employ a truncated regression approach, our inferences based on OLS estimations are unchanged.

TABLE 4
The Association between Alternative Fee Metrics and Discretionary Accruals[a]

$$DCA\_PA = \alpha + \beta_1 FEE + \beta_2 BIG5 + \beta_3 L1ACCRUAL + \beta_4 \ln MVE + \beta_5 MERGER + \beta_6 FINANCING + \beta_7 LEVERAGE + \beta_8 MB + \beta_9 LITIGATION + \beta_{10} INST\_HOLDING + \beta_{11} LOSS + \beta_{12} CFO + \varepsilon$$

Variable	*PADCA*				*REDCA*			
Intercept	−2.432	−2.334	−2.499	−2.255	−2.461	−2.494	−2.664	−2.379
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
FEERATIO	0.270				0.229			
	0.02				*0.04*			
ln *TOTAL*		0.016				−0.015		
		0.59				*0.59*		
ln *AUDIT*			−0.025				−0.058	
			0.47				*0.09*	
ln *NONAUDIT*				0.018				0.007
				0.01				*0.34*
BIG5	−0.227	−0.212	−0.205	−0.226	−0.162	−0.142	−0.138	−0.153
	0.00	*0.01*	*0.01*	*0.00*	*0.04*	*0.07*	*0.08*	*0.05*
L1ACCRUAL	−0.652	−0.658	−0.664	−0.652	−0.743	−0.753	−0.757	−0.747
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
ln *MVE*	−0.073	−0.067	−0.052	−0.073	−0.089	−0.071	−0.060	−0.083
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
MERGER	0.295	0.304	0.311	0.301	0.321	0.334	0.338	0.329
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
FINANCING	0.179	0.186	0.180	0.186	0.221	0.223	0.215	0.226
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
LEVERAGE	−0.532	−0.560	−0.495	−0.561	−0.732	−0.710	−0.642	−0.744
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
MB	0.016	0.016	0.015	0.016	0.014	0.013	0.013	0.014
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
LITIGATION	0.141	0.147	0.145	0.144	0.262	0.266	0.262	0.265
	0.01	*0.01*	*0.01*	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*
INST_HOLDING	−0.548	−0.543	−0.531	−0.557	−0.287	−0.274	−0.263	−0.285
	0.00	*0.00*	*0.00*	*0.00*	*0.01*	*0.02*	*0.02*	*0.01*
LOSS	0.238	0.239	0.246	0.238	0.216	0.223	0.228	0.218
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
CFO	−0.296	−0.304	−0.304	−0.297	−0.358	−0.367	−0.365	−0.364
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Adjusted R^2	0.18	0.18	0.18	0.18	0.21	0.21	0.21	0.21
n	3069	3069	3069	3069	3069	3069	3069	3069

[a] p-values of the estimated parameters are reported in italics.

DCA_PA is the discretionary current accrual measure adjusted for firm performance, which is set equal to *PADCA* or *REDCA*.

Other variables are defined in Table 3.

and both negative *PADCA* and negative *REDCA*. We also find a marginally significant negative relation between *NONAUDIT* and negative *PADCA*. These results are similar to FJN's findings. However, unlike FJN, we find no evidence that *AUDIT* is related to firms' negative discretionary accruals.

Prior research indicates that the U.S. clients of the Big 5 report higher levels of total accruals and smaller amounts of estimated discretionary accruals than the clients of smaller audit firms (Francis et al. 1999). Table 6 reports the results of the discretionary accrual regressions where we partition sample firms based on the sign of their accruals conditioned by their auditor choice. The results of the positive discretionary accrual analyses reported in Panel A indicate that our earlier findings of no relation between positive discretionary accruals and the alternative fee metrics is stable across firms' auditor choice. These results are different from FJN who report a positive relation between *FEERATIO* and positive discretionary accruals reported by non-Big 5 clients. Similar to FJN, however, the results

TABLE 5
Signed Accruals Analysis[a]

Panel A: Positive Accruals

Variable	*+PADCA*				*+REDCA*			
Intercept	−2.136	−2.286	−2.324	−2.011	−2.153	−2.325	−2.414	−2.084
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
FEERATIO	0.008				0.110			
	0.96				*0.48*			
ln *TOTAL*		−0.037				−0.046		
		0.37				*0.24*		
ln *AUDIT*			−0.046				−0.068	
			0.35				*0.16*	
ln *NONAUDIT*				0.015				0.007
				0.11				*0.46*
BIG5	−0.161	−0.150	−0.151	−0.175	−0.155	−0.136	−0.136	−0.156
	0.11	*0.14*	*0.14*	*0.09*	*0.12*	*0.18*	*0.18*	*0.12*
L1ACCRUAL	−0.190	−0.195	−0.192	−0.185	−0.339	−0.354	−0.352	−0.340
	0.08	*0.07*	*0.08*	*0.09*	*0.00*	*0.00*	*0.00*	*0.00*
ln *MVE*	−0.092	−0.075	−0.077	−0.104	−0.120	−0.093	−0.092	−0.119
	0.00	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
MERGER	0.346	0.354	0.353	0.341	0.284	0.300	0.301	0.285
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
FINANCING	0.231	0.225	0.224	0.235	0.264	0.258	0.254	0.266
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
LEVERAGE	−0.483	−0.428	−0.413	−0.505	−0.722	−0.658	−0.621	−0.734
	0.00	*0.01*	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
MB	0.016	0.015	0.015	0.016	0.013	0.011	0.011	0.013
	0.00	*0.00*	*0.00*	*0.00*	*0.01*	*0.02*	*0.02*	*0.01*
LITIGATION	0.150	0.151	0.149	0.147	0.298	0.299	0.297	0.298
	0.04	*0.04*	*0.04*	*0.05*	*0.00*	*0.00*	*0.00*	*0.00*
INST_HOLDING	−0.619	−0.615	−0.613	−0.628	−0.321	−0.316	−0.311	−0.325
	0.00	*0.00*	*0.00*	*0.00*	*0.04*	*0.04*	*0.04*	*0.04*
LOSS	−0.262	−0.256	−0.256	−0.261	−0.209	−0.206	−0.206	−0.208
	0.00	*0.00*	*0.00*	*0.00*	*0.01*	*0.01*	*0.01*	*0.01*
CFO	−0.749	−0.750	−0.747	−0.742	−0.842	−0.847	−0.842	−0.843
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Adjusted R^2	0.17	0.17	0.17	0.17	0.21	0.21	0.21	0.21
n	1551	1551	1551	1551	1629	1629	1629	1629

(*continued on next page*)

TABLE 5 (continued)

Panel B: Negative Accruals

Variable	*−PADCA*				*−REDCA*			
Intercept	2.909	2.641	2.922	2.707	2.879	2.820	3.086	2.817
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
FEERATIO	−0.497				−0.339			
	0.00				*0.04*			
ln *TOTAL*		−0.050				−0.004		
		0.20				*0.93*		
ln *AUDIT*			0.021				0.063	
			0.66				*0.18*	
ln *NONAUDIT*				−0.018				−0.002
				0.08				*0.82*
BIG5	0.205	0.180	0.170	0.186	0.086	0.062	0.057	0.062
	0.10	*0.15*	*0.17*	*0.14*	*0.48*	*0.61*	*0.64*	*0.61*
L1ACCRUAL	0.937	0.941	0.955	0.940	1.084	1.091	1.099	1.091
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
ln *MVE*	0.050	0.049	0.022	0.039	0.057	0.044	0.023	0.044
	0.01	*0.05*	*0.35*	*0.05*	*0.01*	*0.08*	*0.33*	*0.03*
MERGER	−0.186	−0.200	−0.213	−0.205	−0.332	−0.346	−0.351	−0.346
	0.02	*0.01*	*0.01*	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*
FINANCING	−0.086	−0.112	−0.104	−0.107	−0.151	−0.163	−0.151	−0.163
	0.27	*0.15*	*0.19*	*0.17*	*0.06*	*0.04*	*0.06*	*0.04*
LEVERAGE	0.594	0.687	0.569	0.628	0.864	0.871	0.763	0.868
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
MB	−0.016	−0.016	−0.014	−0.016	−0.014	−0.013	−0.012	−0.013
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
LITIGATION	−0.085	−0.096	−0.091	−0.094	−0.142	−0.148	−0.142	−0.148
	0.25	*0.19*	*0.21*	*0.20*	*0.06*	*0.05*	*0.07*	*0.05*
INST_HOLDING	0.345	0.341	0.309	0.345	0.165	0.144	0.117	0.146
	0.03	*0.04*	*0.06*	*0.03*	*0.31*	*0.38*	*0.47*	*0.37*
LOSS	−0.754	−0.754	−0.770	−0.759	−0.685	−0.694	−0.711	−0.694
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
CFO	0.039	0.052	0.055	0.048	0.046	0.058	0.059	0.057
	0.63	*0.51*	*0.49*	*0.55*	*0.58*	*0.48*	*0.47*	*0.48*
Adjusted R^2	0.26	0.25	0.25	0.26	0.27	0.27	0.27	0.27
n	1518	1518	1518	1518	1440	1440	1440	1440

[a] p-values of the estimated parameters are reported in italics.
Variable definitions are provided in Table 3.

of the negative discretionary accrual analyses reported in Panel B indicate that the negative association between *FEERATIO* and negative discretionary accruals documented in the full sample does not hold for non-Big 5 clients.[16] Overall, the results of our discretionary accrual

---

[16] Auditor choice and auditor fees are correlated with firm size. We perform three additional tests to investigate whether size affects our results. First, we replace the fee metrics in the discretionary accrual models with the residuals from the respective fee regressions (Equation (1)). We find that the association between *FEERATIO* residuals and the absolute value of discretionary accruals is positive and significant. However, after partitioning

(*continued on page 630*)

TABLE 6
Big 5 versus Non-Big 5 Discretionary Accrual Analysis[a]

Panel A: Positive Discretionary Accruals

	Big 5								Non-Big 5							
	+PADCA				+REDCA				+PADCA				+REDCA			
FEERATIO	0.012				0.151				−0.178				−0.131			
	0.94				*0.38*				*0.69*				*0.76*			
ln *TOTAL*		−0.015				0.030				−0.157				−0.117		
		0.73				*0.49*				*0.23*				*0.34*		
ln *AUDIT*			−0.012				−0.060				−0.210				−0.034	
			0.82				*0.25*				*0.18*				*0.81*	
ln *NONAUDIT*				0.011				0.001				0.024				0.026
				0.33				*0.96*				*0.16*				*0.10*
Adjusted R^2	0.15	0.15	0.15	0.15	0.21	0.21	0.21	0.21	0.15	0.16	0.16	0.16	0.21	0.21	0.21	0.22
n	1324	1324	1324	1324	1408	1408	1408	1408	227	227	227	227	221	221	221	221

Panel B: Negative Discretionary Accruals

	Big 5								Non-Big 5							
	−PADCA				−REDCA				−PADCA				−REDCA			
FEERATIO	−0.609				−0.394				0.081				0.062			
	0.00				*0.02*				*0.88*				*0.92*			
ln *TOTAL*		−0.067				0.001				−0.118				−0.279		
		0.10				*0.98*				*0.42*				*0.11*		
ln *AUDIT*			0.016				0.078				−0.101				−0.263	
			0.74				*0.11*				*0.51*				*0.15*	
ln *NONAUDIT*				−0.021				−0.004				−0.025				0.002
				0.06				*0.69*				*0.28*				*0.95*
Adjusted R^2	0.26	0.25	0.25	0.26	0.28	0.28	0.28	0.28	0.25	0.25	0.25	0.25	0.21	0.23	0.23	0.21
n	1389	1389	1389	1389	1305	1305	1305	1305	129	129	129	129	135	135	135	135

---

[a] p-values of the estimated parameters are reported in italics.

The results are based on the model presented in Table 4 omitting the *BIG5* variable.

Variable definitions are provided in Table 3.

We do not report the coefficients on the control variables, which, in general, are similar to those reported in Tables 4 and 5.

analyses indicate that the relation between auditor fees and biased financial reporting is sensitive to the estimate of discretionary accruals and the audit fee metric used to capture the economic bond between the auditor and its client.

Klein (2002) and Kothari et al. (2002) state that measurement error in discretionary accruals can lead to erroneous inferences when the measurement error is correlated with the test variable. To investigate whether this issue explains why our *FEERATIO* results differ from FJN's in the positive discretionary accrual analyses, we do the following. First, we calculate *DIFFERENCE* as being equal to *DCA*, i.e., the estimate of discretionary current accruals used by FJN, minus *PADCA*.[17] *DIFFERENCE* represents the bias in the measurement of discretionary current accruals when the estimate of discretionary current accruals does not control for performance. Second, we correlate *DIFFERENCE* with *FEERATIO*, i.e., the test variable of interest, conditional on firms reporting income-increasing or income-decreasing DCA.

We find that for positive (negative) DCA firms *DIFFERENCE* is positively (negatively) correlated with *FEERATIO*. The positive correlation between *DIFFERENCE* and *FEERATIO* for firms reporting income-increasing discretionary current accruals indicates that as *FEERATIO* increases the measurement error in FJN's dependent variable increases. Thus, the measurement bias in the dependent variable likely contributes to FJN's findings of a positive and significant coefficient on *FEERATIO* in their positive *DCA* analysis. For firms reporting income-decreasing accruals, as *FEERATIO* increases the difference between FJN's estimate and our estimate of discretionary current accruals decreases, which likely explains why our results are similar to FJN's results for negative accrual firms.

It appears that FJN's finding of a positive relation between *FEERATIO* and the magnitude of income-increasing accruals is driven by systematic bias in their discretionary accrual measure that is correlated with *FEERATIO*.

Earnings Benchmark Tests

FJN examine the relation between earnings benchmarks and auditor fees to provide further evidence on whether the provision of nonaudit services is associated with earnings management. Specifically, FJN investigate the association between auditor fees and the likelihood of firms reporting small earnings increases (*INCREASE*) and the likelihood of firms meeting or beating analyst earnings forecasts (*SURPRISE*). We estimate the following logit regression to replicate FJN's earnings benchmark results:

$$\begin{aligned} BENCHMARK = {} & \omega_0 + \omega_1 FEE + \omega_2 LITIGATION - \omega_3 MB + \omega_4 \ln MVE \\ & + \omega_5 INST\_HOLDING + \omega_6 LOSS + \omega_7 PACDA + \varepsilon \end{aligned} \qquad (8)$$

where:

Footnote 16, continued

firms based on the sign of their discretionary accruals, we find a negative association between *FEERATIO* residuals and negative discretionary accruals. These results are consistent with our main analysis and suggest that endogeneity is not an issue. Our second test partitions firms into quintiles based on their fiscal year-end total assets. In general, the inferences drawn from the results do not change. Finally, in our last test, we restrict the sample to firms having at least $100 million in total assets. Once again, the inferences drawn from results are the same as our initial analyses. In combination, these three tests suggest that our results are not driven by some residual size effect.

[17] This comparison is based on performance-adjusted discretionary current accruals, whereas FJN's tabled results are based on total discretionary accruals not adjusted for firm performance. FJN claim that their results hold for performance-adjusted current discretionary accruals.

BENCHMARK = *INCREASE* or *SURPRISE*;
INCREASE = 1 when the difference between a firm's 2000 and 1999 net income (Compustat data 172), scaled by beginning of year *MVE*, falls in the interval [0.00, 0.02), and 0 otherwise; and
SURPRISE = 1 when a firm meets or beats by 1 cent the mean consensus analysts forecast, as reported by FirstCall, for fiscal year 2000, and 0 otherwise.

All other variables are as previously defined.

Our benchmark regression model is different than FJN's in that we exclude audit tenure, *CFO*, *MERGER*, *FINANCING*, *ROA*, and annual market adjusted return from our regression model because FJN report insignificant coefficients on these determinants. We add *PADCA* to the *BENCHMARK* regression model to control for firms' discretionary accruals, which likely is one mechanism that firms use to meet earnings benchmarks.[18]

Panel A of Table 7 reports descriptive statistics for the sample used in the *SURPRISE* analysis. Requiring firms to have an analyst following reduces the sample to 1,666 firms, with the average sample firm being relatively larger than the average firm of the full sample. Panel B of Table 7 reports the *SURPRISE* regression results. We find no statistically significant association between any of the audit fee metrics and *SURPRISE*. These results also hold after partitioning firms by their auditor choice. Our results differ from FJN, who find a positive association between *SURPRISE* and *FEERATIO*. FJN also report a positive (negative) association between *RANKNON* (*RANKAUD*) and *SURPRISE*. We are unable to provide an explanation as to why our results differ from those of FJN.

Panel A of Table 8 provides descriptive statistics for the variables used in the *INCREASE* benchmark regression. The descriptive statistics on the explanatory variables are similar to those of the discretionary accrual analysis. Panel B of Table 8 reports the *INCREASE* regression results. We find negative and significant coefficients on *TOTAL* and *AUDIT*, which is consistent with the findings of FJN, but is inconsistent with the claim that greater auditor fees increases the likelihood of biased financial reporting. We also partition sample firms by auditor choice and find similar results.[19]

In summary, the results of FJN earnings management analyses indicate a significant relation between FEERATIO and alternative proxies for earnings management when earnings are biased upward. We find, however, no indication of a significant relation between *FEERATIO* and upward biased earnings management proxies after controlling for firm performance.

Market Reaction Tests

FJN investigate the market reaction to the first-time disclosure of audit and nonaudit fees by calculating abnormal returns (*ARET*) cumulated over one-, two-, and three-day windows centered on the proxy filing date. After controlling for size, FJN report negative associations between *FEERATIO*, *RANKNON*, and *TOTAL*, and one-day *ARET*. We have two primary concerns related to FJN's market reaction analysis. First, when we calculate abnormal returns on firms' proxy filing dates and, as FJN note, we find that the magnitude

---

[18] The sign and significance of the coefficients on the audit fee metrics are the same when we exclude *PADCA* from the benchmark models; however, the explanatory power of the model is lower.

[19] The sample used in the *INCREASE* analysis is equal to the discretionary accrual sample reduced for 121 utility firms and 190 firms that did not have beginning-of-year-2000 market value of equity available on Compustat. In addition, we note no difference in our inferences when we run the earnings benchmark tests excluding loss firms.

TABLE 7
Meeting Earnings Benchmarks: Earnings Surprise

Panel A: Descriptive Statistics (n = 1666)

Variable	Mean	Median
FEERATIO	53.74%	57.06%
TOTAL	2.33	0.61
AUDIT	0.66	0.25
NONAUDIT	1.67	0.32
PADCA	−2.81%	0.51%
MB	3.92	2.22
INST_HOLDING	41.47%	42.36%
BIG5	0.95	1.00
LITIGATION	0.39	0.00
LOSS	0.32	0.00
SURPRISE	0.27	0.00

Panel B: Association Tests[a]

$$SURPRISE = \omega_0 + \omega_1 FEE + \omega_2 LITIGATION + \omega_3 MB + \omega_4 \ln MVE + \omega_5 INST\_HOLDING + \omega_6 LOSS + \omega_7 PACDA + \varepsilon$$

Variable	Full (n = 1666)				BIG 5 (n = 1585)			
intercept	−1.457	−1.468	−1.655	−1.382	−1.542	−1.521	−1.697	−1.413
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
FEERATIO	0.326				0.333			
	0.24				*0.25*			
ln *TOTAL*		−0.026				−0.108		
		0.67				*0.76*		
ln *AUDIT*			−0.076				−0.066	
			0.27				*0.35*	
ln *NONAUDIT*				0.000				0.008
				0.98				*0.73*
LITIGATION	0.232	0.235	0.218	0.243	0.260	0.265	0.247	0.270
	0.06	*0.05*	*0.08*	*0.05*	*0.04*	*0.04*	*0.05*	*0.03*
MB	0.008	0.006	0.005	0.007	0.005	0.004	0.003	0.005
	0.37	*0.51*	*0.57*	*0.45*	*0.55*	*0.69*	*0.76*	*0.61*
ln *MVE*	0.085	0.114	0.131	0.101	0.095	0.121	0.137	0.105
	0.03	*0.01*	*0.00*	*0.01*	*0.01*	*0.01*	*0.00*	*0.01*
INST_HOLDING	−0.241	−0.248	−0.222	−0.256	−0.200	−0.209	−0.185	−0.216
	0.34	*0.32*	*0.38*	*0.31*	*0.43*	*0.41*	*0.47*	*0.40*
LOSS	−0.973	−0.963	−0.971	−0.963	−0.993	−0.984	−0.991	−0.984
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
PADCA	0.818	0.803	0.804	0.806	0.809	0.795	0.795	0.797
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Pseudo R^2	0.05	0.04	0.05	0.04	0.05	0.05	0.05	0.05

---

[a] p-values of the estimated parameters are reported in italics.

SURPRISE = 1 if the firm meets or beats by one cent the mean analysts forecast for fiscal 2000, 0 otherwise. The remaining variables are defined in Table 3.

TABLE 8
Meeting Earnings Benchmarks: Small Earnings Increase

Panel A: Descriptive Statistics (n = 2758)

Variable	Mean	Median
FEERATIO	45.83%	46.95%
TOTAL	1.68	0.39
AUDIT	0.51	0.19
NONAUDIT	1.18	0.17
PADCA	−0.15%	1.10%
MB	3.47	1.72
INST_HOLDING	34.99%	31.07%
BIG5	0.88	1.00
LITIGATION	0.40	0.00
LOSS	0.42	0.00
INCREASE	0.18	0.00

Panel B: Association Tests

$$INCREASE = \omega_0 + \omega_1 FEE + \omega_2 LITIGATION + \omega_3 MB + \omega_4 \ln MVE + \omega_5 INST\_HOLDING + \omega_6 LOSS + \omega_7 PACDA + \varepsilon$$

Variable	Full Sample (n = 2758)				BIG 5 (n = 2421)			
Intercept	−3.199	−4.908	−4.348	−3.346	−3.252	−4.162	−4.477	−3.362
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
FEERATIO	−0.120				−0.036			
	0.66				*0.90*			
ln *TOTAL*		−0.253				−0.258		
		0.00				*0.00*		
ln *AUDIT*			−0.321				−0.343	
			0.00				*0.00*	
ln *NONAUDIT*				−0.019				−0.016
				0.27				*0.41*
LITIGATION	0.461	0.378	0.341	0.455	0.418	0.332	0.288	0.414
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.01*	*0.02*	*0.00*
MB	0.012	0.008	0.009	0.012	0.024	0.016	0.016	0.023
	0.06	*0.23*	*0.20*	*0.07*	*0.01*	*0.06*	*0.06*	*0.01*
ln *MVE*	0.323	0.445	0.440	0.332	0.314	0.448	0.450	0.325
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
INST_HOLDING	0.028	0.103	0.098	0.039	0.080	0.144	0.147	0.091
	0.91	*0.68*	*0.69*	*0.88*	*0.76*	*0.58*	*0.57*	*0.73*
LOSS	−1.333	−1.312	−1.332	−1.331	−1.350	−1.327	−1.351	−1.345
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
PADCA	0.826	0.837	0.843	0.838	1.120	1.122	1.119	1.132
	0.01	*0.01*	*0.00*	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*
Pseudo R^2	0.16	0.17	0.17	0.16	0.16	0.17	0.17	0.16

---

[a] p-values of the estimated parameters are reported in italics.

INCREASE = 1 when the change in net income scaled by MVE $t-1$ is with in the interval [0, .02), and 0 otherwise. The remaining variables are defined in Table 3.

of the abnormal returns is economically insignificant. Panel A of Table 9 reports the mean of one-, two-, and three-day market-adjusted cumulative abnormal returns (CAR). The one- and two-day CARs are less than one half of 1 percent, and neither is significantly different from zero. The three-day CAR is significantly different from zero, but is less than 1 percent of firm value and is positive.

Our second concern is that firms are required to make a substantial number of disclosures in their proxy statements.[20] Consequently, efforts to examine the relation between fee metrics and abnormal returns without controlling for the other information contained in proxy statements can result in incorrect inferences drawn on whether the market reacts to the disclosure of auditor fees. Under the assumption that the economic events driving all proxy disclosures other than the disclosure of auditor fees, on average, are the same from year to year across the market, we calculate a change in market return (*CHCAR*) defined as $CAR_{2000} - CAR_{1999}$. We view this as a reasonable assumption because audit fee data was first disclosed in firms' 2000 proxy statements. Panel B of Table 9 reports the mean values of *CHCAR* for the one-, two-, and three-day windows. None of the values are significant from zero. Panel C of Table 9 reports the results of regressing *CHCAR* on the fee metrics after controlling for firm size. We find a negative association between *CHCAR* and *TOTAL* in the one-day *CHCAR* regression, and a negative association between *CHCAR* and *AUDIT* in the one-day and two-day *CHCAR* regressions. Unlike FJN, we find no association between *FEERATIO* and *CHCAR*, regardless of event window. The inconsistencies between our results and FJN's results suggest that it is difficult to draw reliable conclusions about the market's reaction to audit fee disclosures without controlling for the firm-specific disclosures contained in proxy statements.

V. CONCLUSION

This study investigates FJN's results that indicate that nonaudit fees are positively associated with measures of biased financial reporting.[21] We replicate FJN's empirical tests to investigate whether their results are sensitive to research design choices. Table 10 summarizes our results and how they compare to FJN. The key differences between our results and FJN's results are as follows. First, we find no relation between positive discretionary accruals and any of the auditor fee metrics when discretionary accruals are adjusted for firm performance and sample firms are partitioned by income-increasing versus income-decreasing accruals. Second, in the earnings benchmark tests, we find no relation between fee ratio and the likelihood that firms beat analysts' forecasts. Finally, we find no evidence that the market reacts to the magnitude of nonaudit fees relative to total fees. Our results, along with the results reported in other concurrent audit fee research (e.g., Chung and Kallapur 2003; Francis and Ke 2001; DeFond et al. 2002) do not support FJN's conclusion that firms purchasing nonaudit services manage earnings to a greater extent than other firms.

---

[20] For example, firms are required to disclose facts related to directors and executive officers, compensation of directors and executive officers, compensation plans, authorization or issuance of securities, modification or exchange of securities, financial and other information related to mergers, consolidations, acquisitions, disposition of property, and restatement of accounts (Coopers & Lybrand 1997).

[21] There is one important caveat to both FJN's and our research. A cross-sectional analysis, like the one used in our papers, can only tell us the association between audit fee metrics and proxies for biased financial reporting. This enables us to assess whether relatively higher fees are associated with more biased financial reporting. However, it does not enable us to comment on the overall level of auditor independence.

TABLE 9
The Market's Reaction to Audit Fee Disclosures

Panel A: Market Reaction to 2001 Proxy Statements

	n	Mean	p-value
Window			
3-day	2051	0.34%	0.03
2-day	2051	0.20%	0.15
1-day	2051	0.04%	0.68

Panel B: Change in Market Reaction to 2001 versus 2000 Proxy Statements (CHCAR)

	n	Mean	p-value
Window			
3-day	2051	0.33%	0.16
2-day	2051	0.00%	0.99
1-day	2051	0.10%	0.52

(*continued on next page*)

TABLE 9 (continued)

Panel C: Association between CHCAR and Audit Fee Metrics

$$CHCAR = \alpha + \beta_1 FEE + \beta_2 \ln MVE + \varepsilon$$

Variable	Window 3-day	2-day	1-day	3-day	2-day	1-day	3-day	2-day	1-day	3-day	2-day	1-day
Intercept	0.993	−0.246	0.158	1.045	−0.698	−1.208	0.347	−1.502	−1.631	2.090	0.358	−0.209
	0.14	*0.67*	*0.73*	*0.30*	*0.42*	*0.07*	*0.76*	*0.13*	*0.03*	*0.01*	*0.62*	*0.71*
FEERATIO	1.823	1.555	−0.113									
	0.11	*0.11*	*0.88*									
ln *TOTAL*				−0.102	−0.247	−0.427						
				0.69	*0.25*	*0.01*						
ln *AUDIT*							−0.309	−0.480	−0.535			
							0.29	*0.06*	*0.01*			
ln *NONAUDIT*										0.125	0.049	−0.059
										0.18	*0.54*	*0.34*
ln *MVE*	−0.280	−0.091	−0.001	−0.142	0.099	0.190	−0.082	0.151	0.179	−0.278	−0.049	0.036
	0.02	*0.38*	*0.99*	*0.37*	*0.46*	*0.07*	*0.58*	*0.23*	*0.07*	*0.03*	*0.65*	*0.67*
Adjusted R^2	0.002	0.000	0.001	0.001	0.000	0.002	0.001	0.001	0.003	0.001	0.001	0.001
n	2051	2051	2051	2051	2051	2051	2051	2051	2051	2051	2051	2051

---

[a] p-values of the estimated parameters are reported in italics. The p-value reported in Panel B is for a pairwise t-test. All p-values are two tailed tests.

The sample is all firms having daily returns on CRSP over the proxy release window (day −1 to +1) in the current year *t* and previous year *t*−1, where event window is centered on the proxy filing date. We eliminate firms that file pre-proxy statements in 2001.

CAR = the cumulative abnormal return over the respective event window, where the abnormal return is calculated as the firm return minus the CRSP value-weighted market return; and

$CHCAR = CAR_{it} - CAR_{it-1}$.

All other variables are defined in Table 3.

TABLE 10
Summary of Results

This table summarizes our primary results and the primary results of Frankel et al. (2002) (FJN). The empirical models used in our analyses differ from those used by FJN (see the text for details). The sample observations also differ between the two papers, although both samples are representative of the population of Compustat firms and are drawn from fiscal 2000 proxy statements. *TOTAL* is the sum of audit and nonaudit fees. *FEERATIO* is the ratio of nonaudit fees to total auditor fees. *RANKTOT* is the percentile rank of total fees by auditor.

	Frankel et al. (2002)		Our Results	
Analysis	***FEERATIO***	***RANKTOT***	***FEERATIO***	***TOTAL***
Discretionary Accruals				
Unsigned Discretionary Accruals	Positive	No association	Positive	No association
Positive Discretionary Accruals	Positive	Positive	No association	No association
Negative Discretionary Accruals	Negative	No association	Negative	No association
Positive Discretionary Accruals–non-Big 5	No association	Negative	No association	No association
Earnings Benchmarks				
Analyst Surprise Benchmark	Positive	No association	No association	No association
Earnings Increase Benchmark	No association	Negative	No association	Negative
Event Study				
1-day Market Reaction to Fee Disclosure	Negative	Negative	No association	Negative
3-day Market Reaction to Fee Disclosure	No association	Negative	No association	No association

Stated associations are statistically significant at the <.10 level.

APPENDIX

FJN's Empirical Model

FJN use the following OLS regression model to test the association between the alternative fee metrics and firms' discretionary accruals:

$$\begin{aligned} ABSDACC = {} & \alpha + \beta_1 FEEVAR + \beta_2 BIGFIVE + \beta_3 AUDTEN + \beta_4 CFO \\ & + \beta_5 ABSCFO + \beta_6 ACC + \beta_7 ABSACC + \beta_8 LEVERAGE \\ & + \beta_9 LITRISK + \beta_{10} M/B + \beta_{11} LOGMVE + \beta_{12} \%INST + \beta_{13} LOSS \\ & + \beta_{14} FIN/ACQ + \beta_{15} ANNRET + \varepsilon \end{aligned}$$

where:

FEEVAR = the alternative fee specifications;
BIGFIVE = 1 if the firm hires a Big 5 audit firm, and 0 otherwise;

AUDTEN = 0 to 5 representing the number of years the auditor for fiscal 2000 has audited the firm (Compustat data 149);
CFO = the firm's cash flow from operations (defined as Compustat item 308) scaled by beginning of the year total assets;
ABSCFO = the absolute value of *CFO* scaled by beginning of the year total assets;
ACC = total accruals (defined as net income minus *CFO*) scaled by beginning of the year total assets;
ABSCFO = the absolute value of *ACC* scaled by beginning of the year total assets;
LEVERAGE = the ratio of total liabilities to total assets;
LITRISK = 1 if the firm operates within a high-litigation industry and 0 otherwise, where high-litigation industries are industries with SIC codes of 2833–2836, 3570–3577, 3600–3674, 5200–5961, 7370–7374;
M/B = the firm's market-to-book ratio;
LOGMVE = the log of the firm's market value of equity;
%INST = the percentage of shares held by institutional owners at the beginning of the 2000 calendar year;
LOSS = 1 if the firm reports a net loss in fiscal 2000, and 0 otherwise;
FIN/ACQ = 1 if the firm issued securities or made an acquisition during 2000, and 0 otherwise; and
ANNRET = the percentage compounded monthly return for 2000 adjusted for the CRSP value-weighted market index.

REFERENCES

Beck, P. J., T. J. Frecka, and I. Solomon. 1988. An empirical analysis of the relationship between MAS involvement and auditor tenure: Implications for auditor independence. *Journal of Accounting Literature* 7: 65–84.

Becker, C. L., M. L. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 14: 1–24.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24: 99–126.

Chung, H., and S. Kallapur. 2003. Client importance, nonaudit services, and abnormal accruals. Working paper, Purdue University.

Coopers & Lybrand. 1997. *SEC Manual.* Seventh edition. Somerset, NJ: John Wiley & Sons.

DeAngelo, L. 1981. Auditor independence, "Low Balling," and disclosure regulation. *Journal of Accounting and Economics* (August): 113–127.

Dechow, P., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70: 193–225.

DeFond, M., K. Raghunandan, and K. Subramanyam. 2002. Do nonaudit service fees impair auditor independence? Evidence from going concern audit opinions. *Journal of Accounting Research* 40: 1247–1274.

Firth, M. 1997. The provision of nonaudit services by accounting firms to their audit clients. *Contemporary Accounting Research* 14 (Summer): 1–21.

Francis, J., E. L. Maydew, and H. C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18: 17–34.

———, and B. Ke. 2001. Do nonaudit services compromise auditor independence? Working paper, University of Missouri and Pennsylvania State University.

Frankel, R., M. Johnson, and K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* 77: 71–105.

Kasznik, R. 1999. On the association between voluntary disclosure and earnings management. *Journal of Accounting Research* 37: 57–81.

Klein, A. 2002. Audit committee, board characteristics, and earnings management. *Journal of Accounting and Economics* 33: 375–400.

Kothari, S., A. Leone, and C. Wasley. 2002. Performance matched discretionary accrual measures. Working paper, University of Rochester.

Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, New York, NY, September 28.

———. 2000. Renewing the covenant with investors. Remarks delivered at NYU Center for Law and Business, New York, NY, May 10.

Magee, R., and Tseng. 1990. Audit pricing and independence. *The Accounting Review* 65 (April): 315–336.

Myers, J., L. Myers, and T. Omer. 2003. Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *The Accounting Review* (July): 779–799.

Read, W., and S. Tomczyk. 1992. An examination of changes in scope of services performed by CPA firms. *Accounting Horizons*: 42–51.

Securities and Exchange Commission (SEC). 1978. *Disclosure of Relationships with Independent Public Accountants*. Accounting Series Release No. 250. June 29. Washington, D.C.: Government Printing Office.

Simunic, D. 1984. Auditing, consulting, and auditor independence. *Journal of Accounting Research* 22: 679–702.

Warfield, T. D., J. J. Wild, and K. L. Wild. 1995. Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics* 20: 61–91.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 641–678

The World Price of Earnings Opacity

Utpal Bhattacharya
Indiana University

Hazem Daouk
Cornell University

Michael Welker
Queen's University

ABSTRACT: We analyze financial statements from 34 countries for the period 1984–1998 to construct a panel data set measuring three dimensions of reported accounting earnings for each country: earnings aggressiveness, loss avoidance, and earnings smoothing. We hypothesize that these three dimensions are associated with uninformative or opaque earnings, and so we combine these three measures to obtain an overall earnings opacity time-series measure per country. We then explore whether our three measures of earnings opacity affect two characteristics of an equity market in a country: the return the shareholders demand and how much they trade. While not all results are consistent for our three individual earnings opacity measures, our panel data tests document that, after controlling for other influences, an increase in overall earnings opacity in a country is linked to an economically significant increase in the cost of equity and an economically significant decrease in trading in the stock market of that country.

Keywords: *earnings opacity; earnings aggressiveness; loss avoidance; earnings smoothing; cost of equity; trading volume.*

Data Availability: *The data used in this study are publicly available from the sources indicated in the text.*

JEL Classification: *G15, G30, M41.*

We are grateful for the discussions on earnings opacity with Rajesh Aggarwal, George Allayanis, Daniel Beneish, Walt Blacconiere, Alex Butler, Art Durnev, Mara Faccio, Andreas Hackethal, Randy Heron, Paul Hribar, Ben Jacobsen, Christian Leuz, Ross Levine, Jamie Pratt, Scott Richardson, Avanidhar Subrahmanyam, Dan Thornton, Jim Wahlen, and Kelley Welch, and for comments from an anonymous referee, the editor, and workshop participants at The University of Alabama, University of Amsterdam, Cornell University, University of Houston, Indiana University, University of Kansas, University of Ljubljana, University of Maryland, University of Michigan, New York University, Queen's University, University of Padua, University of Western Ontario, Wilfrid Laurier University, the 2002 Georgia Tech International Finance Conference, the 2002 Canadian Academic Accounting Association Meeting, the 2002 European Finance Association Meeting, the 2002 SFM Conference in Taiwan, and the 2002 CIFRA Conference in Chile. We thank Joon Ho Hwang and Sue Zheng for being such careful research assistants. Professor Bhattacharya thanks Queen's University for research support when he was a Resident Scholar there during Summer 2001. An earlier version of this paper was titled "The World Price of Earnings Management."

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted May 2002
Accepted December 2002

I. INTRODUCTION

The recent decline in equity values in the United States has been attributed to investor concerns over corporate governance and accounting.[1] This decline is consistent with investors perceiving a decreased informativeness in U.S. accounting numbers, and demanding higher rates of return to compensate for this additional component of informational risk.[2] The purpose of our paper is to investigate whether informational risk associated with accounting earnings impacts equity markets around the world. Of necessity, we do this using joint tests. We first measure distributional properties of accounting earnings that suggest poor correspondence between observable accounting earnings and unobservable economic earnings—which we dub "earnings opacity"—in a country every year, and we then determine whether our measures of earnings opacity in a country are associated with the return shareholders demand for holding equity in that country and are associated with shareholder trading of equity in that country.[3]

We define the earnings opacity of a country as the extent to which the distribution of reported earnings of firms in that country fails to provide information about the distribution of the true, but unobservable, economic earnings of firms in that country. As reported earnings of a particular firm in a country equals unobservable economic earnings plus a noise term, earnings opacity of a country is simply the average lack of informativeness of reported earnings in that country.[4]

Reported earnings in a country could be opaque because of a complex interaction among, at least, three factors: managerial motivation, accounting standards, and the enforcement of accounting standards (e.g., audit quality). It could be that earnings are opaque because managers are motivated to manipulate earnings, and they can do this either because accounting standards allow substantial flexibility, accounting standards do not exist to specify accounting principles related to some areas of business activity, or accounting standards, though rigorous, are weakly enforced. It could also be that earnings are opaque, not because managers manipulate earnings, but simply because accounting standards do not call for accounting treatments that transparently reflect underlying business activity and management is not willing or able to overcome these deficiencies by voluntarily providing more informative earnings reports.

Earnings opacity is inherently difficult to measure, particularly across countries, because it is not possible to pinpoint management's motives, and it is difficult to compare accounting standards and the enforcement of these accounting standards. In addition, it is not possible to capture all factors that might influence earnings opacity, or to model how they interact to produce more or less opaque earnings. So, instead of studying the inputs that determine earnings opacity, we analyze the outcome: the distributional properties of reported accounting numbers across countries and across time that suggest earnings opacity. Specifically, we use measures that are intended to capture three attributes of earnings numbers that could lead to earnings opacity: earnings aggressiveness, loss avoidance, and earnings smoothing. We focus on these three dimensions because the literature has suggested that these three dimensions may weaken the link between accounting performance and the true economic performance of a firm. We limit our analysis to industrial firms, so that differences in the underlying earnings process across different industry groups, and differences in the proportion of firms in various industry groups across countries and across time, do not affect the dimensions of reported earnings we examine. Finally, given the above mentioned difficulties in measuring earnings opacity, all our tests are inherently joint tests of two hypotheses: one, our three measures, or a composite of all three, are associated with uninformative or opaque earnings; and two, earnings opacity creates an informational risk that affects the cost of equity and trading in the stock market.

---

[1] A direct consequence of this attribution is the Sarbanes-Oxley Act of 2002, which was signed into law on July 30, 2002. One of the main objectives of this law is "to protect investors by improving the accuracy and reliability of corporate disclosures."

[2] Information risk refers to a variety of risks that investors may face as a result of possessing inadequate or imprecise information on which to base their investment decisions.

[3] In a previous version of this paper, we had also investigated the effect of earnings opacity in a country on U.S. equity holdings in that country. Because of lack of data, our tests were cross-sectional and not panel data tests. As the number of countries were roughly of the same order of magnitude as the number of control variables, these cross-sectional tests had embarrassingly few degrees of freedom. We dropped this section.

[4] Our definition respects the goals of financial reporting laid out in various statements of the Financial Accounting Standards Board (FASB). One such statement reads, "The primary focus of financial reporting is information about an enterprise's performance provided by measures of earnings and its components." (See FASB 1978, SFAC No. 1, para. 43).

We construct a panel data set for each of these three measures of earnings opacity—earnings aggressiveness, loss avoidance, and earnings smoothing—and then combine them to obtain an overall earnings opacity time-series measure per country. We find that our estimates of average earnings opacity per country are significantly associated with variables that might impact the overall quality of a financial reporting regime of a country, namely the Center for International Financial Analysis and Research (CIFAR) disclosure index and the number of auditors per 100,000 population.

The second part of our paper goes on to investigate whether earnings opacity affects equity markets. We first examine the effect of our measures of earnings opacity on the return shareholders demand for holding equity (cost of equity). We measure the effect on the cost of equity employing two distinct approaches used in a companion paper (Bhattacharya and Daouk 2002). We discuss the details of these approaches, and their merits and demerits, in the next section of the paper.

The first approach is to extract the cost of equity from the dividend discount model. After controlling for other influences, we find in our panel data tests that the earnings aggressiveness dimension and overall earnings opacity have significant adverse effects on the cost of equity. An increase in our measure of overall earnings opacity from the 25th percentile rank to the 75th percentile rank is associated with a 2.8 percent increase in the cost of equity measured using dividend yields.

The second approach uses an international asset pricing factor model. This empirical specification allows for partial integration of a country to the world equity markets. After controlling for other influences, we find in our panel data tests that the loss avoidance dimension and overall earnings opacity have a significant effect on the cost of equity. An increase in our measure of overall earnings opacity from the 25th percentile rank to the 75th percentile rank is associated with a 3.2 percent increase in the cost of equity measured using this model.

Our last set of panel data tests examines the effect of earnings opacity on the level of trading. The details of the data set used to measure trade are discussed in the next section. After controlling for other influences, we find that earnings aggressiveness, earnings smoothing, and overall earnings opacity have significant adverse effects on trade. An increase in our measure of overall earnings opacity from the 25th percentile rank to the 75th percentile rank is associated with an 8.8 percent decrease in annual trade.

To summarize, we find that, after controlling for other influences, an increase in our measure of overall earnings opacity in a country is linked to an increase in the cost of equity and a decrease in trading in the stock market of that country. Both these effects are economically as well as statistically significant. However, our tests also reveal some inconsistencies in the relationship between our individual measures of earnings opacity and our cost of equity and trade measures. As all our tests are joint tests, this suggests that caution must be exercised in interpreting the links between our individual measures based on reported earnings and earnings opacity, as well as the links between these individual measures and the cost of equity and trade.

The rest of the paper is organized as follows. Section II provides the conceptual development underlying our empirical analyses. Section III discusses the methodological issues in the measurement of the earnings opacity variables as well as the stock market variables—cost of equity and trade. In Section IV we discuss the data and give some summary statistics. Section V contains the results of the main empirical tests. We conclude by discussing the implications and limitations of our analysis in Section VI.

II. CONCEPTUAL DEVELOPMENT

Bushman and Smith (2001), who call for more research using cross-country designs to explore the links between financial accounting information and corporate governance, identify three channels by which earnings opacity may affect financial markets. First, better accounting information helps investors distinguish between good and bad investments, which decreases estimation risk, which decreases the firm's cost of equity. Second, better accounting information helps investors distinguish between good and bad managers, which decreases agency costs, which decreases the firm's cost of equity. Third, earnings opacity, by weakening the link between reported accounting earnings and unobservable economic earnings, increases asymmetric information. An increase in asymmetric information leads to an increase in the adverse selection problem a liquidity provider faces when trading with insiders. The liquidity providers in such a market would protect themselves by increasing their sell price and decreasing their buy price.[5] This increases the transaction cost, which induces a shareholder to require an even higher return on equity for compensation,[6] and also leads shareholders to trade less often or not trade at all.[7]

Our research investigates whether the three properties of reported accounting earnings we examine result in more opaque earnings, leading to increased information risk that is associated with increased cost of equity and decreased trading. Four assumptions underlie our above hypothesis. First, our measures actually measure what we claim they measure; that is, our measures of the three properties of reported accounting earnings are positively associated with how opaque, or uninformative, investors perceive reported accounting earnings to be. Second, markets are efficient in the sense that investors can detect the level of earnings opacity, but they cannot "see through" it. Third, the informational asymmetry created by earnings opacity is not completely resolved through some other communication mechanism, like alternate disclosures directed at large, affiliated stakeholders. Fourth, the informational risk caused by earnings opacity is an important factor relative to the other factors that affect equity markets, and so it is priced. None of these assumptions may hold. The above assumptions are tested in the second part of this paper.

A cross-country comparison of earnings opacity has many advantages over a cross-firm comparison of earnings opacity. First, because of considerable differences in accounting standards and audit quality across the globe, we can obtain an enviable dispersion in earnings opacity around the world. Second, as Bushman and Smith (2001) state, the cross-country differences in earnings opacity can be linked meaningfully to the cross-country differences in economic efficiency and institutional factors. A cross-country design also has some disadvantages. Two potential disadvantages—a more severe omitted variables problem and a more severe endogeneity problem—are discussed later in the paper.

---

[5] See Glosten and Milgrom (1985) and Kyle (1985) for formal models.

[6] See Amihud and Mendelson (1986) for a formal model on why this should happen for riskless assets. Jacoby et al. (2000) extend this to risky assets. Brennan and Subrahmanyam (1996) provide convincing empirical evidence.

[7] See Bhattacharya and Spiegel (1991) for an analysis of the critical level of asymmetric information needed for a market breakdown.

Ours is not the first paper to exploit the advantages of cross-country comparisons. Our paper is a part of the growing international accounting literature that examines the value-relevance of accounting measures (Alford et al. 1993; Ali and Hwang 2000; Harris et al. 1994; Joos and Lang 1994; Land and Lang 2002). This value-relevance literature examines associations between economic income, measured using equity returns, and accounting data such as earnings. This literature reveals which countries experience greater associations between accounting data and equity returns, but it does not examine the potential consequences of variation in earnings informativeness on equity markets. Recent literature has examined international variation in analyst forecasts (Ashbaugh and Pincus 2001; Chang et al. 2000), but this literature also does not examine equity market consequences of this variation. Finally, examinations of earnings timeliness and conservatism (Ball, Kothari, and Robin 2000), or the effect of institutional factors on earnings management (Leuz et al. 2002), explore the potential causes of variation in earnings informativeness around the world, but again do not address the effects of this variation on equity markets. Our contribution to the above literature is that ours is the first paper, to our knowledge, that measures earnings opacity at a country level every year to form a panel data set and then uses panel data tests to check whether earnings opacity adversely affects the equity markets of that country. Our paper should be viewed as complementary to Leuz et al. (2002), who measure earnings management at a cross-sectional level across 31 countries and then explore whether institutional factors are linked to the cross-sectional differences in earnings management. Our paper should also be viewed as complementary to a recent survey conducted by PricewaterhouseCoopers (2001) that constructs a broad measure of opacity in a particular country, and links it to capital inflows and the country risk premium in sovereign bonds of that country.

III. VARIABLE SELECTION AND MEASUREMENT ISSUES

Earnings Opacity Measures

Earnings Aggressiveness Measure

Our first measure of earnings opacity is earnings aggressiveness. Ball, Kothari, and Robin (2000) argue that the opposite of aggressiveness, accounting conservatism, which is the more timely incorporation of economic losses versus economic gains into accounting earnings, arises to reduce information asymmetry. Specifically, they argue that three factors are expected to lead to accounting conservatism. First, accountants are aware that managers would like to report economic gains and suppress information about economic losses. Hence, accountants find negative information more credible, and are more likely to incorporate it into accounting income. Second, lenders are important users of financial statements, and lenders are more impacted by economic losses than by economic gains. Third, the timely incorporation of economic losses provides an important corporate governance role, providing quick feedback about bad investment decisions and strategies that managers may not wish to disclose. The first and third of these factors suggest that accounting conservatism is related to informativeness, since conservative accounting is expected to provide information that management might have incentives to withhold otherwise.

It is possible that earnings aggressiveness does not necessarily lead to earnings opacity. It could be argued that conservative accounting prevents good news from being transmitted quickly, thus adding noise. However, given that one might reasonably expect managerial incentives to overstate rather than understate earnings on average, our belief is that aggressive earnings are more opaque earnings, because such accounting reports are more likely to reflect biased and optimistic reporting on the part of management, adding noise to reported earnings and, hence, increasing earnings opacity. To understand these managerial motives, see, for example, Rangan (1998), Teoh et al. (1998), Shivakumar (2000), Healy (1985), Barth et al. (1999). Ultimately, whether or not earnings aggressiveness leads to earnings opacity is an empirical issue.

We use accruals to measure earnings aggressiveness. As earnings aggressiveness is the tendency to delay the recognition of losses and speed the recognition of gains, it implies that if cash flow realizations are held equal, then we would expect accruals to increase as earnings aggressiveness increases. For example, lower of cost or market rules, accounting procedures with a conservative bias, result in negative accruals. Aggressive accounting would be characterized by fewer such negative accruals that capture economic losses, and more positive accruals that capture economic gains, increasing the overall level of accruals.[8] Though it is true that unrealized gains and unrealized losses would eventually be recognized in accounting earnings in any clean surplus accounting system, the more conservative accounting system is expected to result in more negative accruals at any given point in time because a greater proportion of economic losses relative to economic gains will be reflected in accounting earnings at any point in time. This motivates us to measure earnings aggressiveness of a country at a point in time as the median for country i, year t, of accruals divided by lagged total assets. We use the median observation of scaled accruals to minimize the influence of extreme observations. The higher the median observation of scaled accruals of country i in year t, the higher the earnings aggressiveness. The effect of earnings aggressiveness on the distribution of accounting earnings vis-à-vis economic earnings is depicted in the Earnings Aggressiveness graph of Figure 1.[9]

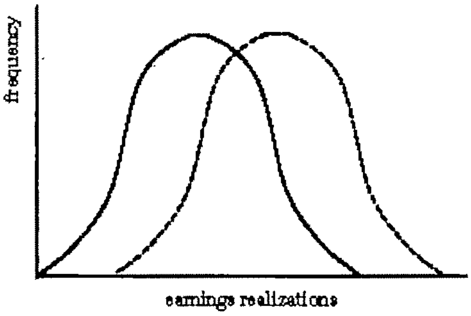
Consistent with much of the past literature (e.g., Dechow et al. 1995; Healy 1985; Jones 1991; Leuz et al. 2002), we compute scaled accruals from balance sheet and income statement information, and then compute scaled cash flows as scaled operating income minus scaled accruals. We do not use information from the cash flow statement because of differences in the presentation of cash flow information across countries and time. In fact,

---

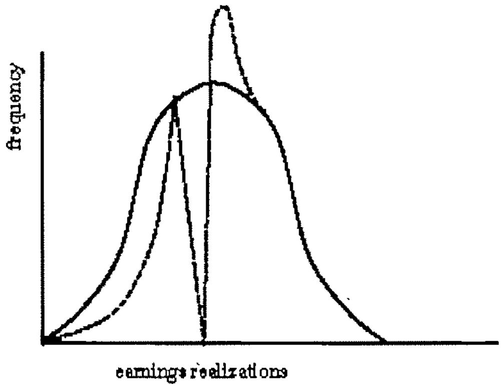
[8] For example, aggressive revenue recognition may result in an increase in accounts receivable, thus increasing accruals.

[9] Ball, Kothari, and Robin (2000), following Basu (1997), have used an alternative way to measure conservatism, which is to check whether negative economic income, as reflected in negative security returns, is more quickly incorporated in accounting earnings than positive economic income. However, this metric is inappropriate for our research design because we are interested in examining the effects of earnings opacity on equity market variables, and using equity market variables to measure earnings opacity would introduce circularity. Ball, Robin, and Wu (2000) also adopt an alternative metric from Basu (1997), which is based on the time-series properties of accounting earnings. Unfortunately, this time-series approach would use all available time-series data to generate one measure of accounting conservatism per country, and thus cannot be used to generate a panel data set. Hence, the level of accruals is used as our proxy for earnings aggressiveness. We do, however, acknowledge that accruals are a noisy measure of earnings aggressiveness; accruals measure accounting aggressiveness with error and measure other facets of accounting earnings as well.

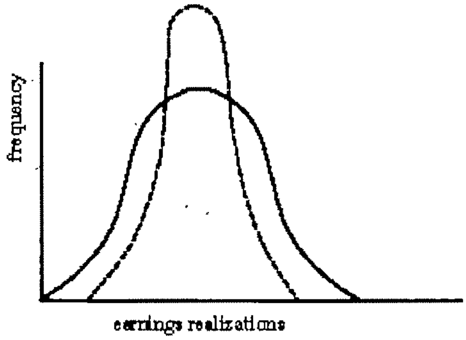
FIGURE 1
Distributional Properties of Accounting Earnings versus Economic Earnings



Earnings Aggressiveness



Loss Avoidance



Earnings Smoothing

——— Economic earnings
- - - - Accounting earnings

many of our sample countries do not require the preparation or presentation of a statement of cash flows.[10] We define scaled accruals as:

$$ACC_{kt} = (\Delta CA_{kt} - \Delta CL_{kt} - \Delta CASH_{kt} + \Delta STD_{kt} - DEP_{kt} + \Delta TP_{kt})/TA_{kt-1} \tag{1}$$

where:

ACC_{kt} = scaled accruals for firm k, year t;
ΔCA_{kt} = change in total current assets for firm k, year t;
ΔCL_{kt} = change in total current liabilities for firm k, year t;
$\Delta CASH_{kt}$ = change in cash for firm k, year t;
ΔSTD_{kt} = change in current portion of long-term debt included in total current liabilities for firm k, year t;
DEP_{kt} = depreciation and amortization expense for firm k, year t;
ΔTP_{kt} = change in income taxes payable for firm k, year t; and
TA_{kt-1} = total assets for firm k, year $t-1$.

Loss Avoidance Measure

Our second measure of earnings opacity is loss avoidance behavior. Burgstahler and Dichev (1997) present persuasive evidence that U.S. firms engage in earnings management to avoid reporting negative earnings. Degeorge et al. (1999) provide evidence that suggests that the following hierarchy exists among three earnings thresholds: (1) avoiding negative earnings, (2) reporting increases in quarterly earnings, and (3) meeting analysts' earnings forecasts. As Burgstahler and Dichev (1997) and Degeorge et al. (1999) discuss, these results indicate that incentives to report positive earnings (i.e., beat a benchmark of zero earnings) exist for some sample firms. Such loss avoidance behavior obscures the relationship between earnings and economic performance, thus increasing earnings opacity.

We define firms with small positive earnings (small negative earnings) as firms with net income scaled by lagged total assets between 0 and 1 percent (between 0 and -1 percent). We find the ratio of the number of firms with small positive earnings minus the number of firms with small negative earnings divided by their sum. The higher is this ratio in country i, year t, the higher is the loss avoidance. The effect of loss avoidance on the distribution of accounting earnings vis-à-vis economic earnings is depicted in the Loss Avoidance graph of Figure 1.

Earnings Smoothing Measure

Our third measure of earnings opacity is earnings smoothing. Some accounting standards (for example, cases of high book/tax conformity) or some managerial motives (see, for example, Trueman and Titman 1988; Fudenberg and Tirole 1995) may lead to smooth earnings over time. If accounting earnings are artificially smooth, then they fail to depict the true swings in underlying firm performance, thus decreasing the informativeness of reported earnings and, hence, increasing earnings opacity. This is consistent with the view of earnings smoothing taken in Leuz et al. (2002). An alternative view, as expressed in Zarowin (2002), is that earnings smoothing can be used by management as a means to convey information, potentially decreasing earnings opacity. While we believe that earnings smoothing at the country level is indicative of accounting that obscures information about economic volatility, whether or not earnings smoothing leads to earnings opacity and adverse capital market consequences is again an empirical issue.

---

[10] We repeat all our tests with three modifications to the above definitions of accruals and cash flow. One modification dropped the subtraction of depreciation and amortization from the definition of accruals. This specification focuses on just working capital accruals. A second modification replaced operating income with net income in the definition of cash flow, and dropped the add back of the change in taxes payable from the definition of accruals. This specification includes taxes in the definition of income and accruals. Finally, we included the change in total reserves and the change in deferred taxes and dropped the add back of the change in taxes payable from our definition of accruals, and replaced operating income with net income in the definition of cash flows. This follows the definition of accruals and cash flows used in Ali and Hwang (2000), and it includes the changes in deferred taxes and the reserve accounts that exist in some countries in our definition of accruals. Under all three modifications, the results of this paper are qualitatively unchanged.

Following Leuz et al. (2002), we find the cross-sectional correlation between the change in accruals and the change in cash flows, both scaled by lagged total assets, in country i, year t. Cash flows are obtained by subtracting accruals (which were obtained in Equation (1)) from operating earnings. Because some degree of earnings smoothing is a natural outcome of any accrual accounting process, this measure is expected to be negative on average. However, the more negative this correlation, the more likely it is that earnings smoothing is obscuring the variability in underlying economic performance, and the greater is the earnings opacity. The effect of earnings smoothing on the distribution of accounting earnings vis-à-vis economic earnings is depicted in the Earnings Smoothing graph of Figure 1.

Overall Earnings Opacity Measure

We rank all the raw time-series earnings aggressiveness median observations, across countries and years, into deciles, with higher ranks associated with greater earnings aggressiveness; we rank all the raw time-series loss avoidance ratios, across countries and years, into deciles, with higher ranks associated with greater loss avoidance; we rank all the raw time-series earnings smoothing correlations, across countries and years, into deciles, with higher ranks associated with greater earnings smoothing. Hence, each dimension of earnings opacity in a country each year is assigned a rank between 1 and 10, depending on which decile of the earnings opacity dimension distribution across *all* country-years that particular country for that year appears. For example, a rank of 3 for Australia in 1986 in the earnings smoothing dimension means that Australia in 1986 was in the third decile of all earnings smoothing measures across all countries and across all years. We then average the earnings aggressiveness rank, the loss avoidance rank, and the earnings smoothing rank in each country-year to obtain a time-series of overall earnings opacity for each country.

While there is no strong conceptual basis for aggregation into an overall earnings opacity measure, we perform the aggregation for two reasons. First, to the extent that each of our three earnings opacity dimensions measures the same underlying phenomenon with (uncorrelated) measurement error, our aggregate measure will suffer less from measurement error than each of the individual measures. Second, to the extent that each of our three earnings opacity dimensions measures a unique dimension of earnings opacity, then, intuitively, each separate dimension adds to the overall difficulty investors have in gleaning information from reported accounting earnings. We report all our results for each separate dimension of earnings opacity as well as for our aggregate measure.

To construct raw or rank cross-sectional measures of each individual dimension of earnings opacity per country, we average across all available years for each country the raw or the rank measure, respectively, of each dimension of earnings opacity. To construct rank cross-sectional measures of overall earnings opacity per country, we average across all available years the overall earnings opacity rank for each country.

Limitations of Our Earnings Opacity Measures

There are limitations of our measures of earnings opacity. First, we want to know how expected earnings opacity affects investor behavior, but expected earnings opacity is unobservable. We use lagged earnings opacity measures from year $t-1$ in all our tests to proxy for investor expectations about earnings opacity in year t, which implicitly assumes that investors observe earnings opacity after the fact, and their expectation of earnings opacity this year are based on their observation of last year's earnings opacity.

Second, it is possible that our measures of the distributional characteristics of reported earnings are impacted by factors that do not affect earnings opacity. For example, industry membership and growth may systematically impact the distributional characteristics of accruals, cash flows, and earnings, and these effects may be clear and predictable to investors, resulting in no effect on their perception of earnings opacity. As previously mentioned, we control for industry by limiting the sample to industrial firms. We control for growth by including the real GDP growth rates in all our empirical tests. These, however, may be over-controls if the variation in the distributional characteristics of reported earnings caused by industry membership and growth do indeed affect earnings opacity.

Third, while we have argued that earnings aggressiveness, loss avoidance, and earnings smoothing increase earnings opacity, counter arguments, mentioned above, also exist. Clear evidence linking our measures to informativeness in accounting numbers does not exist. Hence, our tests are tests of two joint hypotheses: our measures are correlated positively with earnings opacity, and earnings opacity affects cost of equity and trading in stock markets. Insignificant results could be due to a lack of a link between earnings opacity and cost of equity and trading, or it could be that our measures do not capture earnings opacity.

Stock Market Measures

Cost of Equity Measures

The cost of equity in a country is defined as the return shareholders require for holding shares in that country. This is an expectations variable, which we measure using *ex post* data. We employ two approaches used in a companion paper (Bhattacharya and Daouk 2002).

The first approach is to compute the cost of equity by backing it out from the classical constant growth dividend discount model. Bekaert and Harvey (2000, Appendix A) explore in great detail the relationship between dividend yields and the cost of equity for many models. Assuming that the best forecast for future growth rates in dividends is the most current dividend growth rate, which implies that we assume that dividend growth rates follow a random walk, it follows that the estimated cost of equity = current dividend yield × (1 + current growth rate of dividends) + current growth rate of dividends.

The advantages of using dividend yields to measure cost of equity are many. Dividend yields are observable, stable, and stationary. A sharp change in cost of equity should lead to a sharp change in dividend yields. The disadvantage of using dividend yields is that changes in dividend yields may come about because of repurchases of stock, and may come about because of changes in growth opportunities. The first factor is not much of a problem in emerging markets because repurchases are minor. We try to control for the second factor—growth opportunities—by including GDP growth rates as control variables.

If the earnings opacity variables have no incremental effect on the cost of equity, then those variables will be orthogonal to the above estimate of the cost of equity. We control for other influences.

The second approach to estimating the cost of equity accounts explicitly for risk. The international version of the capital asset pricing model does not hold up well in the data (see Harvey [1991] or Ferson and Harvey [1993]). The consensus seems to be that a country's beta with respect to the world market portfolio has some merit to explain expected returns for developed countries; the variance of return of the country's stock market does better in explaining expected returns for emerging markets (see Harvey 1995).

We adopt a simplified version of Bekaert and Harvey (1995) as our international asset pricing model. Their empirical specification allows for partial integration of a country to the world equity markets. Their model is very appealing because it permits a country to evolve from a developing segmented market (where risk is measured by the country's variance) to a developed country that is integrated to world equity markets (where risk is measured by the sensitivity of a country's equity returns to movements in the world market portfolio). The special case of complete integration, where the world factor is the only factor, is nested in their model. This international asset pricing model is expressed as follows:

$$(r_{i,t}) - (r_{f,t}) = \alpha_0 + \phi_{i,t}\lambda_{cov}h_{i,w,t} + (1 - \phi_{i,t})\lambda_{var}h_{i,t} + e_{i,t} \tag{2}$$

where:

$r_{i,t}$ = the dollar monthly return of the stock market index of country i at time t;
$r_{f,t}$ = the monthly return of the one-month U.S. T-Bill at time t;
α_0 = a constant that would be estimated;
$\phi_{i,t}$ = a measure of the level of integration of country i at time t, $0 < \phi_{i,t} < 1$;
λ_{cov} = the price of the covariance risk that would be estimated;
$h_{i,w,t}$ = the conditional covariance of the monthly return of the stock market index of country i with the monthly return of the world index at time t;
λ_{var} = the price of own country variance risk that would be estimated (which we are restricting to be the same across all countries);
$h_{i,t}$ = the conditional variance of the monthly return of the stock market index of country i at time t; and
$e_{i,t}$ = the residual error term.

The independent variables in Equation (2)—conditional covariance $h_{i,w,t}$ and conditional variance $h_{i,t}$—are separately estimated pair-wise for each country i and world pair from the multivariate ARCH model specified below:

$$\begin{aligned}
r_{i,t} &= c_1 + \varepsilon_{i,t}, \\
r_{w,t} &= c_2 + \varepsilon_{w,t}, \\
h_{i,t} &= b_1 + a_1\left(\frac{1}{2}\varepsilon^2_{i,t-1} + \frac{1}{3}\varepsilon^2_{i,t-2} + \frac{1}{6}\varepsilon^2_{i,t-3}\right), \\
h_{w,t} &= b_2 + a_2\left(\frac{1}{2}\varepsilon^2_{w,t-1} + \frac{1}{3}\varepsilon^2_{w,t-2} + \frac{1}{6}\varepsilon^2_{w,t-3}\right), \\
h_{i,w,t} &= b_3 + a_3\left(\frac{1}{2}\varepsilon_{i,t-1}\varepsilon_{w,t-1} + \frac{1}{3}\varepsilon_{i,t-2}\varepsilon_{w,t-2} + \frac{1}{6}\varepsilon_{i,t-3}\varepsilon_{w,t-3}\right), \\
\varepsilon_{i,t}, \varepsilon_{w,t} &\sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}h_{i,t} & h_{i,w,t}\\ h_{i,w,t} & h_{w,t}\end{bmatrix}\right).
\end{aligned} \tag{3}$$

where:

$r_{w,t}$ = the dollar monthly return of the stock market index of the world at time t;
$\varepsilon_{i,t-j}$ = the innovation in monthly return of the stock market index of country i at time $t-j$, $j \in \{0,1,2,3\}$;
$\varepsilon_{w,t-j}$ = the innovation in monthly return of the stock market index of the world at time $t-j$, $j \in \{0,1,2,3\}$; and
$h_{w,t}$ = the conditional variance of the monthly return of the stock market index of the world at time t.

Equation (3) was first introduced by Bollerslev et al. (1988). As in Engle et al. (1987), the weights of the lagged residual vectors are taken to be 1/2, 1/3, and 1/6, respectively. The constants a_2, b_2, and c_2 are constrained to be identical for all country-world pairs. Maximum likelihood is used to estimate Equation (3).

The other independent variable in Equation (2)—$\phi_{i,t}$—measures the level of integration of country i at time t. We define it as follows:

$$\phi_{i,t} = \frac{\exp\left(a_1\left(\frac{exports_{i,t} + import_{i,t}}{gdp_{i,t}}\right)\right)}{1 + \exp\left(a_1\left(\frac{exports_{i,t} + imports_{i,t}}{gdp_{i,t}}\right)\right)}. \tag{4}$$

The definition of $\phi_{i,t}$ in Equation (4) implies that it is a function of the ratio of the sum of *exports* and *imports* to *gross domestic product*. It is designed to take on values between 0 and 1. When its value is 0, the country is not integrated with world equity markets, and its equity is exposed only to local risk (own variance). When its value is 1, the country is fully integrated with world equity markets, and its equity is exposed only to global risk (covariance with world factor). Bekaert and Harvey (1997) find that increases in this ratio are empirically associated with increased importance of the world factor relative to local risk factors.

We use a two-step procedure (first remove the effect of risk, and then test the effect on residuals) instead of using a one-step procedure (include all independent variables in Equation (2) directly). We do so because of technical convergence problems in the one-step nonlinear estimation procedure. If the earnings opacity variables have no incremental effect on the cost of equity, then those variables will be orthogonal to the residuals from the model in Equation (2). We control for other influences on this residual. The advantage of using a well-specified asset pricing factor model like Equation (2) to measure cost of equity is that we explicitly account for risk. This comes at a price. Recall that all the independent variables in Equation (2) are estimates from other models. This introduces estimation error, which may introduce bias, and it reduces power.

Trade Measure

A good metric to capture the amount of trade in a market is turnover, which is defined as the ratio of volume of dollar trade per month to dollar market capitalization at the end of the month. To mitigate the effect of outliers, which occur because the denominator is small in some countries, we take the natural logarithm of this ratio.

IV. DATA AND DESCRIPTIVE STATISTICS

Earnings Opacity Measures

The data used in constructing the earnings opacity variables come from the Worldscope database for the years 1984 through 1998. We restrict the sample to industrial firms (SIC codes 2000–3999 and SIC codes 5000–5999) to increase the homogeneity of our sample across countries and across time. Since the underlying earnings process being represented by accounting earnings is similar for industrial firms, this restriction reduces the probability that the cross-country differences and time-differences we observe in our earnings opacity measures are caused by the difference in or changes in industrial composition in our sample. This sample restriction is also consistent with much of the accounting literature (e.g., Alford et al. 1993; Ali and Hwang 2000).[11] Because all our tests are panel data tests, we include countries that have data for more than three years, and have more than 20 firms per year. This yields 58,653 firm-year observations from 34 countries, spanning the years 1986 through 1998. (We lose 1984 and 1985 because the calculation of changes in accruals and cash flows requires data from years $t-1$ and $t-2$.)

The names of the countries for which we have data are given in Column (1) in the Appendix, the sample period for each country is given in Column (2), and the number of firm-years per country is given in Column (3). For each firm-year, we use the following variables from Worldscope: cash, total current assets, total current liabilities, income taxes payable, current portion of long-term debt included in total current liabilities, depreciation and amortization expense, operating income, net income, and total assets. Some firms do not have information on income taxes payable or on the current portion of long-term debt included in total current liabilities. Similar to Leuz et al. (2002), if these variables are missing, then we assume them to be zero. We include observations with fiscal years ending between July 1 of year t and June 30 of year $t+1$ in the calculation of our earnings opacity variables for year t. So, for example, observations with fiscal years ending between July 1, 1995 and June 30, 1996 are considered year 1995 observations.

Descriptive information on each of the raw earnings opacity variables for each sample country is provided in Columns (2) through (4) of Table 1. Each column gives the average across the available years for each country for each measure. Column (2) provides the average accruals divided by lagged total assets for our sample countries. As expected, average accruals are negative, averaging about 2 percent of lagged total assets. Interestingly, three of the 34 countries in our sample—Greece, India, and Turkey—have positive accruals. The loss avoidance measure is presented in Column (3). Avoidance of small negative bottom-line earnings is observed in 32 of our 34 countries, implying that this is a global phenomenon. Finally, the earnings smoothing measure—the average cross-sectional correlation between the change in cash flows and the change in accruals—is presented in Column (4). As expected, the correlation is strongly negative in every country in our sample.

Other Financial Reporting Measures

We first analyze the relation between our cross-sectional measures of earnings opacity and alternative cross-country measures related to financial reporting quality that have been documented in the past literature. As discussed previously, we expect earnings opacity to

---

[11] We also ran all our tests using a broader sample consisting of all nonfinancial firms (i.e., we excluded only SIC codes in the 6000s) in a previous version of this paper. Such a sample has been constructed by Leuz et al. (2002) and Land and Lang (2002). Inferences from this expanded sample are qualitatively similar to our reported results.

TABLE 1
Summary Statistics

	Earnings Opacity			Other Variables								
(1) Countries	(2) Earnings Aggressiveness	(3) Loss Avoidance	(4) Earnings Smoothing	(5) Auditors per 100,000 Population	(6) Disclosure Level	(7) IAS Use	(8) Common Law	(9) Mean Monthly Return/ Standard Deviation of Monthly Return	(10) Trade	(11) Real % GDP Growth (1985–1998)	(12) Insider Trading Enforcement Date	(13) Liberalization Date
Australia	−0.0213	−0.04615	−0.82374	539	80	0	Yes	0.1292	−3.3862	3.14	1996	Before 01/86
Austria	−0.03727	0.500397	−0.87909	NA	62	0	No	0.0980	−3.5173	2.49	No	Before 01/86
Belgium	−0.05467	0.317765	−0.87866	38	68	0	No	0.3282	−4.6080	2.21	1994	Before 01/86
Brazil	−0.0068	0.035416	−0.77614	1	NA	1	No	0.0285	NA	2.70	1978	05/91
Canada	−0.03433	0.450318	−0.81781	350	75	0	Yes	0.1446	−3.6756	2.54	1976	Before 01/86
Chile	−0.01706	0.6	−0.91368	87	78	NA	No	0.2628	−4.9177	6.75	1996	01/92
Denmark	−0.03937	0.267444	−0.91274	106	75	0	No	0.2198	−5.0497	2.37	1996	Before 01/86
Finland	−0.03267	0.621092	−0.88223	10	83	0	No	0.1572	NA	2.12	1993	Before 01/86
France	−0.03827	0.376352	−0.86549	45	78	1	No	0.2164	−4.0970	2.13	1975	Before 01/86
Germany	−0.04138	0.586525	−0.8978	26	67	0	No	0.1671	−1.9795	2.96	1995	Before 01/86
Greece	0.01344	0.652206	−0.91468	12	61	NA	No	0.1708	−4.0892	2.15	1996	12/87
Hong Kong	−0.01194	0.17013	−0.85786	110	73	0	Yes	0.1455	−3.401	5.30	1994	Before 01/86
India	0.001681	0.735644	−0.86787	9	61	1	Yes	0.0301	−3.7879	5.73	1998	11/92
Indonesia	−0.00098	0.733766	−0.85613	2	NA	NA	No	−0.0876	−4.5513	4.55	1996	09/89
Ireland	−0.024	0.153846	−0.86847	262	81	1	Yes	0.2213	NA	5.95	No	Before 01/86
Italy	−0.02733	0.505334	−0.92531	110	66	0	No	0.1339	−4.2832	1.93	1996	Before 01/86
Japan	−0.01247	0.642863	−0.92135	10	71	0	No	0.0698	−4.2534	2.74	1990	Before 01/86
Korea (South)	−0.0115	0.595265	−0.93793	7	68	0	No	0.0433	−2.9382	7.48	1988	12/88
Malaysia	−0.01226	0.469553	−0.87234	48	79	2	Yes	0.0271	−4.3673	6.50	1996	05/89
Mexico	−0.02058	−0.03333	−0.74486	15	71	1	No	0.1243	−3.3118	3.27	No	Before 01/86
The Netherlands	−0.04506	0.378023	−0.9172	52	74	1	No	0.3426	−2.8995	2.74	1994	Before 01/86
Norway	−0.03786	0.178788	−0.72913	NA	75	1	No	0.0839	−3.4892	2.94	1990	02/91
Pakistan	−0.02584	0.616327	−0.91133	2	73	2	Yes	0.0369	NA	4.26	No	07/86
Portugal	−0.06614	0.211112	−0.87479	NA	NA	1	No	0.0822	−5.1253	4.38	No	Before 01/86
Singapore	−0.02534	0.484873	−0.88578	273	79	1	Yes	0.0958	−4.0609	6.97	1978	01/92

(continued on next page)

TABLE 1 (continued)

	Earnings Opacity			Other Variables								
(1) Countries	(2) Earnings Aggressiveness	(3) Loss Avoidance	(4) Earnings Smoothing	(5) Auditors per 100,000 Population	(6) Disclosure Level	(7) IAS Use	(8) Common Law	(9) Mean Monthly Return/ Standard Deviation of Monthly Return	(10) Trade	(11) Real % GDP Growth (1985–1998)	(12) Insider Trading Enforcement Date	(13) Liberalization Date
South Africa	−0.02021	0.307692	−0.88157	35	79	1	No	0.0609	−4.2894	1.30	No	Before 01/86
Spain	−0.0379	0.514142	−0.85582	18	72	0	No	0.2321	−3.3038	3.18	1998	Before 01/86
Sweden	−0.02256	0.340096	−0.84528	41	83	0	No	0.2242	−3.6422	1.49	1990	Before 01/86
Switzerland	−0.03963	0.589985	−0.87921	53	80	1	No	0.2602	−3.3903	1.51	1995	Before 01/86
Taiwan	−0.02405	0.691198	−0.85276	17	58	NA	No	0.0950	−2.3109	7.26	1989	01/91
Thailand	−0.03953	0.730403	−0.85693	5	66	2	Yes	0.0542	−3.3376	5.97	1993	09/87
Turkey	0.127142	0.5	−0.67144	NA	58	1	No	0.0698	−3.6251	4.82	1996	08/89
United Kingdom	−0.02924	0.372985	−0.8683	352	85	0	Yes	0.2369	−3.1222	2.57	1981	Before 01/86
United States	−0.03833	0.350638	−0.77688	168	76	0	Yes	0.3097	−2.7766	2.73	1961	Before 01/86
All Countries	−0.02141	0.38765	−0.86541	93	73			0.1416	−3.7196	3.74		

(1) Annual financial statement data for firms in 20 developed markets and 14 emerging markets were obtained from Worldscope. These countries are listed in Column (1).

(2) We scale accruals by lagged total assets for each firm, determine its median in the cross-section of firms per country per year, and then average across time to obtain the "earnings aggressiveness" variable per country. This is listed in Column (2).

(3) We define firms with small positive earnings (small negative earnings) as firms with net income scaled by lagged total assets between 0 and 1% (between 0 and −1%). We subtract the number of firms with small negative earnings from the number of firms with small positive earnings per country per year, divide this difference by the sum of the two, and then average this ratio across time to obtain the "loss avoidance" variable per country. This is listed in Column (3).

(4) We find the correlation between the change in accruals and the change in operating cash flows (both scaled by lagged total assets) in the cross-section of firms per country per year, and then average across time to obtain the "earnings smoothing" variable per country. This is listed in Column (4).

(5) The number of auditors per 100,000 population in Column (5) comes from Saudagaran and Diga (1997, Table 6). The original source is the International Federation of Accountants (IFAC) Secretariat, 8/13/1996.

(6) Disclosure level data in Column (6) come from Saudagaran and Diga (1997, Table 2). The original source is the Center for International Financial Analysis and Research (CIFAR 1995). The higher the number, more is the disclosure.

(*continued on next page*)

TABLE 1 (continued)

(7) International Accounting Standards (IAS) use data in Column (7) come from Choi et al. (1999, exhibit 8.6). They took it from International Accounting Standards Committee (IASC *Insight* dated October, 1997.) 0 = completely independent standard setting, no use of IAS except possibly a comparison with IAS; 1 = separate accounting standards that are based on and similar to IAS in most cases, however, some standards provide more or less choice; 2 = IAS are used as national standards with some modification for local conditions, standards not covered by IAS added.

(8) The common law data in Column (8) come from the CIA World Factbook (2001), http://www.odci.gov/cia/publications/factbook.

(9) Data on monthly stock market indices for the 20 developed markets were obtained from Morgan Stanley Capital Market International (MSCI). Data on monthly stock market indices for the 14 emerging markets were obtained from the International Financial Corporation (IFC). The mean return scaled by the standard deviation of returns is given in Column (9). The sample periods used to calculate these statistics are given in Column (4) of the Appendix.

(10) Trade is defined as the natural logarithm of the ratio of volume of dollar trade per month to dollar market capitalization at the end of the month. Monthly data on these two variables were obtained from Datastream. Their sample periods are given in Columns (5) and (6) in the Appendix. The mean trade per country is given in Column (10).

(11) The real growth in Gross Domestic Product per country per year is obtained from the World Bank, http://www.worldbank.org/research/growth/GDNdata.htm. The average per country is given in Column (11).

(12) The insider trading enforcement date in Column (12) comes from Bhattacharya and Daouk (2002, Table 1).

(13) The official liberalization date, which was obtained from Bekaert and Harvey (2000), is given in Column (13).

Table 4 presents the results from this panel time-series regression. The coefficient of the overall earnings opacity measure (Model (4)) is positive and statistically significant at the 5 percent level. A detailed look at Models (1), (2), and (3) reveals that this significance is coming from the earnings aggressiveness variable, although the coefficients on the other earnings opacity variables have the right sign. This is consistent with our joint hypotheses that our earnings aggressiveness measure and our overall measure are correlated with earnings opacity, and earnings opacity adversely affects the cost of equity. The association is also economically significant. An increase in overall earnings opacity from the 25th percentile rank to the 75th percentile rank is associated with a 2.8 percent increase in the cost of equity.[16] The coefficient on the insider trading enforcement variable has the right sign and is statistically significant, implying that insider trading enforcement causes the cost of equity to drop as seen in Bhattacharya and Daouk (2002). The liberalization indicator has the correct negative sign, but is not significantly different from zero at conventional levels. GDP growth rates are positively related to this measure of the cost of equity.

Using an International Asset Pricing Model

We estimate Equation (2) using nonlinear least squares. The regressions use data for our 34 countries from December 1986 to December 1998 (some countries do not have data for the full time period). The results are given in Panel A of Table 5.

Panel A of Table 5 reveals that though covariance risk seems to have a positive price (λ_{cov} is positive), the estimates are statistically significant only at the 11 percent level. It also reveals that though own country variance risk has a positive price (λ_{var} is positive), the estimates are statistically significant only at the 12 percent level. These results contrast with the results of Bhattacharya and Daouk (2002), who use the same estimation technique and obtain statistical significance, but that is because their estimation was carried out for a longer 1969–1998 sample period.

Using the residuals from Equation (2) as the dependent variable, we run four panel time-series regressions with country-fixed effects. Models (1), (2), (3), and (4) are as previously defined. We correct for country-specific heteroscedasticity and country-specific autocorrelation in each case. We control for liberalization, insider trading enforcement, and GDP growth as before. We control for two other sources of risk that have been documented in the literature—foreign exchange risk (Dumas and Solnik 1995; Ferson and Harvey 1993)[17] as well as liquidity risk (Brennan and Subrahmanyam 1996)[18]—and that change in our sample period.

Panel B of Table 5 presents the results from this panel time-series regression. The coefficient of the overall earnings opacity measure (Model (4)) is positive and statistically significant at the 8 percent level. A detailed look at Models (1), (2), and (3) reveals that

---

[16] This is calculated as 0.001244 (per month) × 12 months × (6.538 (rank of 75th percentile) – 4.692 (rank of 25th percentile)).

[17] As purchasing power parity is not observed in the data, standard models control for a foreign exchange factor (FX factor). This is why we include it. However, because of convergence problems, our estimation is a two-step procedure. Therefore, unlike the standard models, in the first step we strip out the effects of the local variance factor and the world factor and, in the second step in order to isolate the effect of earnings opacity, we strip out the effects of other factors like the FX factor. The FX factor that we use is the conditional covariance of the return of the stock market index of the country with the return a U.S. investor would get if she held the foreign currency. This conditional covariance is obtained by using the multivariate ARCH model we previously discussed in Equation (3)—just replace the world portfolio (w) by the foreign exchange portfolio (ifx).

[18] The proxy for liquidity risk is turnover. Turnover is the ratio of volume of trade to market capitalization. We take the natural logarithm of this ratio for reasons mentioned before.

TABLE 4
Effect of Earnings Opacity on the Cost of Equity (Using Dividend Yields)[a]

$$\text{MODEL: } Cost\ of\ Equity_{i,t} = \beta_0 + \beta_1\ Dimension\ of\ Earnings\ Opacity_{i,t-1} + \beta_2\ Liberalization_{i,t} + \beta_3\ Insider\ Trading\ Enforcement_{i,t} + \beta_4\ GDP\ Growth_{i,t} + u_{i,t}$$

Dependent Variable[b]	Cost of Equity			
Independent Variables[c]	(1)	(2)	(3)	(4)
Earnings aggressiveness[d]	0.0012 (0.0000)			
Loss avoidance[e]		0.0000 (0.9479)		
Earnings smoothing[f]			0.0001 (0.7443)	
Overall earnings opacity[g]				0.0012 (0.0011)
Liberalization [h]	−0.0045 (0.0862)	−0.0040 (0.1301)	−0.0040 (0.1293)	−0.0043 (0.1033)
Insider trading enforcement[i]	−0.0038 (0.0345)	−0.0039 (0.0300)	−0.0039 (0.0320)	−0.0040 (0.0260)
GDP growth[j]	0.0011 (0.0000)	0.0013 (0.0000)	0.0013 (0.0000)	0.0012 (0.0000)

---

[a] The numbers below are coefficient estimates from panel regressions, and are corrected for country-fixed effects, country-specific heteroscedasticity and country-specific autocorrelation. p-values are in parentheses.

[b] The dependent variable "Cost of equity" is constructed as follows. The cost of equity, if backed out from the dividend discount model, equals current dividend yield (1 + expected dividend growth rate) + expected dividend growth rate. Assuming that dividend growth rates follow a random walk, we replace expected values with contemporary values. The dividend yield data were obtained from Datastream for the main stock market of each country. The sample periods for which these data were obtained are given in the Appendix.

[c] The first four independent variables are the earnings opacity variables, whereas the other independent variables are the control variables. The earnings opacity variables are rank variables. A higher rank implies more earnings opacity. The data to construct the earnings opacity variables come from Worldscope.

[d] We scale accruals by lagged total assets for each firm, determine its median in the cross-section of firms per country per year, and then rank these medians across years and across countries. This rank is the "earnings aggressiveness" time-series variable per country.

[e] We define firms with small positive earnings (small negative earnings) as firms with net income scaled by lagged total assets between 0 and 1 percent (between 0 and −1 percent). We subtract the number of firms with small negative earnings from the number of firms with small positive earnings per country per year, divide this difference by the sum of the two, and then rank this ratio across years and across countries. This rank is the "loss avoidance" time-series variable per country.

[f] We find the correlation between the change in accruals and the change in operating cash flows (both scaled by lagged total assets) in the cross-section of firms per country per year, and then rank these correlations across years and across countries. This rank is the "earnings smoothing" time-series variable per country.

[g] The "overall earnings opacity" time-series variable per country is the average of the "earnings aggressiveness" time-series variable per country, the "loss avoidance" time-series variable per country and the "earnings smoothing" time-series variable per country.

[h] The control variable "Liberalization" is an indicator variable. It changes from 0 to 1 in the month after the official liberalization date that was obtained from Bekaert and Harvey (2000). These liberalization dates are given in Table 1.

[i] The control variable "insider trading enforcement" is an indicator variable. It changes from 0 to 1 in the year after the first enforcement of insider trading laws. This date was obtained from Bhattacharya and Daouk (2002). These insider trading enforcement dates are given in Table 1.

[j] The control variable "GDP growth" is the growth rate of gross domestic product of a country every year. These data come from the World Bank, http://www.worldbank.org/research/growth/GDNdata.htm.

this significance is coming from the loss avoidance variable, in contrast with the results from the dividend yield model where the significance was primarily driven by the earnings aggressiveness variable. This is consistent with our joint hypotheses that our loss avoidance measure and our overall measure are correlated with earnings opacity, and earnings opacity adversely affects the cost of equity. Insider trading enforcement and GDP growth are insignificantly related to the cost of equity in this specification, though both variables have the same sign as in the dividend yield specification. The liberalization indicator again has a negative sign, and is statistically significant in this specification. An increase in overall earnings opacity from the 25th percentile rank to the 75th percentile rank is associated with

---

TABLE 5
Effect of Earnings Opacity on the Cost of Equity (Using an International Asset Pricing Factor Model)

Equation 1:

The international asset pricing factor model used for risk-adjusting is:

$$(r_{i,t} - r_{f,t}) = \alpha_0 + \phi_{i,t}\lambda_{cov}h_{i,w,t} + (1 - \phi_{i,t})\lambda_{var}h_{i,t} + e_{i,t}$$

where the measure of integration of country i at time t, $N_{i,t}$, is defined as:

$$\phi_{i,t} = \frac{\exp\left(\alpha_1\left(\dfrac{exports_{i,t} + imports_{i,t}}{gdp_{i,t}}\right)\right)}{1 + \exp\left(\alpha_1\left(\dfrac{exports_{i,t} + imports_{i,t}}{gdp_{i,t}}\right)\right)}$$

and λ_{cov} is the price of the covariance risk with the world, and λ_{var} is the price of own country variance risk. The independent variables are the conditional covariances and variances, $h_{i,w,t}$ and $h_{i,t}$, respectively, and these are obtained from the multivariate ARCH model below:

$$\begin{aligned}
r_{i,t} &= c_1 + \varepsilon_{i,t},\\
r_{w,t} &= c_2 + \varepsilon_{w,t},\\
h_{i,t} &= b_1 + a_1\left(\frac{1}{2}\varepsilon^2_{i,t-1} + \frac{1}{3}\varepsilon^2_{i,t-2} + \frac{1}{6}\varepsilon^2_{i,t-3}\right),\\
h_{w,t} &= b_2 + a_2\left(\frac{1}{2}\varepsilon^2_{w,t-1} + \frac{1}{3}\varepsilon^2_{w,t-2} + \frac{1}{6}\varepsilon^2_{w,t-3}\right),\\
h_{i,w,t} &= b_3 + a_3\left(\frac{1}{2}\varepsilon_{i,t-1}\varepsilon_{w,t-1} + \frac{1}{3}\varepsilon_{i,t-2}\varepsilon_{w,t-2} + \frac{1}{6}\varepsilon_{i,t-3}\varepsilon_{w,t-3}\right),\\
\varepsilon_{i,t},\varepsilon_{w,t} &\sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}h_{i,t} & h_{i,w,t}\\ h_{i,w,t} & h_{w,t}\end{bmatrix}\right).
\end{aligned}$$

where:

$\varepsilon_{i,t-j}$ = the innovation in monthly return of the stock market index of country i at time $t-j$, $j \in \{0,1,2,3\}$; and
$\varepsilon_{w,t-j}$ = the innovation in monthly return of the stock market index of the world at time $t-j$, $j \in \{0,1,2,3\}$.

(*continued on next page*)

TABLE 5 (continued)

Panel A: Some Coefficients of the Risk-Adjustment Model, Equation 1[a]

Dependent Variable[b]	**Excess return of country**
Some Independent Variables[c]	
Covariance of the country's equity return with the world equity return multiplied by the measure of the country's integration with the world	$\lambda_{cov} = 2.1625$ (0.1076)
Variance of the country's equity return multiplied by 1 minus the measure of the country's integration with the world	$\lambda_{var} = 2.6432$ (0.1244)

---

[a] The numbers below are coefficient estimates from the panel regressions described above. p-values are in parentheses.

[b] The dependent variable is the monthly equity return for each country minus the one-month U.S. T-Bill return. The equity return for each country is computed from its stock market index. Data on monthly stock market indices for the 20 developed markets were obtained from Morgan Stanley Capital Market International (MSCI). Data on monthly stock market indices for the 14 emerging markets were obtained from the International Financial Corporation (IFC). The sample periods are given in the Appendix. The data for the one-month U.S. Treasury bill return was obtained from Datastream.

[c] The measure of a country's integration with the world, as defined above, is computed from its exports, imports, and GDP. It is Equation (4) in the text. Data on quarterly/annual GDP, monthly exports, and monthly imports were from the International Financial Statistics of the International Monetary Fund. The statistics for Taiwan come from Datastream. The sample periods are given in the Appendix.

The conditional covariance of the return of the stock market index with the depreciation of the *i*th foreign currency with respect to the dollar at time *t*, defined as the foreign exchange risk and denoted as $h_{i,ifx,t}$, is estimated from the multivariate ARCH model below.

$$
\begin{aligned}
r_{i,t} &= f_1 + \varepsilon_{i,t},\\
r_{ifx,t} &= f_2 + \varepsilon_{ifx,t},\\
h_{i,t} &= e_1 + d_1\left(\frac{1}{2}\varepsilon^2_{i,t-1} + \frac{1}{3}\varepsilon^2_{i,t-2} + \frac{1}{6}\varepsilon^2_{i,t-3}\right),\\
h_{ifx,t} &= e_2 + d_2\left(\frac{1}{2}\varepsilon^2_{ifx,t-1} + \frac{1}{3}\varepsilon^2_{ifx,t-2} + \frac{1}{6}\varepsilon^2_{ifx,t-3}\right),\\
h_{i,ifx,t} &= \varepsilon_3 + d_3\left(\frac{1}{2}\varepsilon_{i,t-1}\varepsilon_{ifx,t-1} + \frac{1}{3}\varepsilon_{i,t-2}\varepsilon_{ifx,t-2} + \frac{1}{6}\varepsilon_{i,t-3}\varepsilon_{ifx,t-3}\right),\\
\varepsilon_{i,t},\varepsilon_{ifx,t} &\sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}h_{i,t} & h_{i,ifx,t}\\ h_{i,ifx,t} & h_{ifx,t}\end{bmatrix}\right).
\end{aligned}
$$

where:

$\varepsilon_{i,t-j}$ = the innovation in monthly return of the stock market index of country *i* at time $t-j$, $j \in \{0,1,2,3\}$; and

$\varepsilon_{ifx,t-j}$ = the innovation in monthly depreciation of the *i*th foreign currency with respect to the dollar at time $t-j$, $j \in \{0,1,2,3\}$.

Equation 2: Residual from Equation 1, $e_{i,t} = \beta_0 + \beta_1$ *Dimension of Earnings Opacity*$_{i,t-1}$
$+ \beta_2$ *Foreign Exchange Risk*$_{i,t}$
$+ \beta_3$ *Liquidity*$_{i,t}$ $+ \beta_4$ *Liberalization*$_{i,t}$
$+ \beta_5$ *Insider Trading Enforcement*$_{i,t}$
$+ \beta_6$ *GDP Growth*$_{i,t}$ $+ v_{i,t}$

(*continued on next page*)

TABLE 5 (continued)

Panel B: Coefficients of Equation 2[a]

Dependent Variable[b]	**Residual from Risk Adjustment Model**			
Independent Variables[c]	**(1)**	**(2)**	**(3)**	**(4)**
Earnings aggressiveness[d]	−0.0001			
	(0.8121)			
Loss avoidance[e]		0.0013		
		(0.0022)		
Earnings smoothing[f]			0.0000	
			(0.9899)	
Overall earnings opacity[g]				0.0014
				(0.0783)
Foreign exchange risk, $h_{i,ifx,t}$[h]	7.4670	7.5572	7.4821	7.6567
	(0.0002)	(0.0002)	(0.0002)	(0.0001)
Liquidity[i]	0.0035	0.0034	0.0035	0.0041
	(0.0190)	(0.0142)	(0.0137)	(0.0042)
Liberalization[j]	−0.0143	−0.0142	−0.0143	−0.0145
	(0.0275)	(0.0029)	(0.0027)	(0.0024)
Insider trading enforcement[k]	−0.0032	−0.0045	−0.0033	−0.0040
	(0.2613)	(0.1220)	(0.2473)	(0.1720)
GDP growth[l]	0.0007	0.0006	0.0007	0.0005
	(0.1920)	(0.2971)	(0.2020)	(0.3852)

---

[a] The numbers below are coefficient estimates from panel regressions, and are corrected for country fixed-effects, country-specific heteroscedasticity and country-specific autocorrelation. p-values are in parentheses.

[b] The dependent variable is the residual from Equation 1.

[c] The first four independent variables are the earnings opacity variables, whereas the other independent variables are the control variables. The earnings opacity variables are rank variables. A higher rank implies more earnings opacity. The data to construct the earnings opacity variables come from Worldscope.

[d] We scale accruals by lagged total assets for each firm, determine its median in the cross-section of firms per country per year, and then rank these medians across years and across countries. This rank is the "earnings aggressiveness" time-series variable per country.

[e] We define firms with small positive earnings (small negative earnings) as firms with net income scaled by lagged total assets between 0 and 1 percent (between 0 and −1 percent). We subtract the number of firms with small negative earnings from the number of firms with small positive earnings per country per year, divide this difference by the sum of the two, and then rank this ratio across years and across countries. This rank is the "loss avoidance" time-series variable per country.

[f] We find the correlation between the change in accruals and the change in operating cash flows (both scaled by lagged total assets) in the cross-section of firms per country per year, and then rank these correlations across years and across countries. This rank is the "earnings smoothing" time-series variable per country.

[g] The "overall earnings opacity" time-series variable per country is the average of the "earnings aggressiveness" time-series variable per country, the "loss avoidance" time-series variable per country and the "earnings smoothing" time-series variable per country.

[h] The control variable "foreign exchange risk" is estimated from the multivariate ARCH model given above.

[i] The control variable "liquidity" is defined as the natural logarithm of the ratio of volume of dollar trade per month to dollar market capitalization at the end of the month. These data were obtained from Datastream for the main stock market of each country. The sample periods for which these data were available are given in the Appendix.

[j] The control variable "Liberalization" is an indicator variable. It changes from 0 to 1 in the month after the official liberalization date that was obtained from Bekaert and Harvey (2000). These liberalization dates are given in Table 1.

[k] The control variable "insider trading enforcement" is an indicator variable. It changes from 0 to 1 in the year after the first enforcement of insider trading laws. This date was obtained from Bhattacharya and Daouk (2002). These insider trading enforcement dates are given in Table 1.

[l] The control variable "GDP growth" is the growth rate of gross domestic product of a country every year. These data come from the World Bank, http://www.worldbank.org/research/growth/GDNdata.htm.

a 3.2 percent increase in the cost of equity.[19] It should be noted here that the point estimates obtained by two completely different methods of estimating the cost of equity—the dividend yield method (implicitly controls for risk, but has less estimation risk) and an international asset pricing model (explicitly controls for risk, but has more estimation risk)—are very similar: 2.8 percent and 3.2 percent, respectively.

Trading

The measure of trade is turnover, which is defined as the ratio of volume of trade to market capitalization. Using the natural logarithm of this ratio as the dependent variable, we run four panel time-series regressions with country-fixed effects. Models (1), (2), (3), and (4) are as previously defined. We correct for country-specific heteroscedasticity and country-specific autocorrelation in each case. We control for liberalization, insider trading enforcement, and GDP growth as before. The panel regressions use data for the 30 countries for which we have trading data from January 1986 to December 1998 (some countries do not have data for the full time period).

Table 6 presents the results from this panel time-series regression. Except for Model (2) whose coefficient is insignificant, the coefficients of the individual earnings opacity measures (Models (1) and (3)) and the coefficient of the overall earnings opacity measure (Model (4)) are negative and statistically significant at the 5 percent level. This result is consistent with our joint hypotheses that our earnings aggressiveness measure, our earnings smoothing measure and our overall measure are correlated with earnings opacity, and earnings opacity adversely affects trading volume. An increase in overall earnings opacity from the 25th percentile rank to the 75th percentile rank is associated with an 8.8 percent decrease in annual trade. The coefficients on liberalization and insider trading enforcement are significant, and have the right sign.

Further Robustness Checks

Despite two features of our research design that we believe help mitigate concerns over omitted explanatory variables and endogeneity—panel data tests and use of lagged opacity variables—we run two tests to further allay concerns.[20] First, we reran all our tests including lagged equity market development, measured as stock market capitalization divided by GDP, as an additional control variable. This variable is measured on an annual basis and comes from Datastream. While there is no direct conceptual basis for including this variable in equations explaining either cost of equity or trade, we include this measure of equity market development as a proxy for unobserved institutional factors that might promote the development of equity markets within a country and thus impact both earnings opacity and our stock market variables. We find that equity market development is unrelated to either of our measures of the cost of equity capital, but it is positively related to trade. Our previous inferences on the effect of earnings opacity variables on cost of equity or trade are not affected by the addition of this control variable.

Second, we estimated a Vector Auto Regression (VAR) model proposed by Sims (1980) in order to explicitly model earnings opacity, the cost of equity, and trade as endogenously determined dependent variables. The endogenous variables are modeled as linear functions of lagged endogenous variables and all exogenous variables in the system. The system of

---

[19] This is calculated as 0.001430 (per month) × 12 months × (6.538 (rank of 75th percentile) – 4.692 (rank of 25th percentile)).

[20] We do not present the results of these additional tests in tables. Interested readers may obtain these tables directly from the authors.

TABLE 6
Effect of Earnings Opacity on Trade[a]

$$\text{Model: } Trade_{i,t} = \beta_0 + \beta_1 \, Dimension\ of\ Earnings\ Opacity_{i,t-1} + \beta_2 \, Liberalization_{i,t} + \beta_3 \, Insider\ Trading\ Enforcement_{i,t} + \beta_4 \, GDP\ Growth_{i,t} + u_{i,t}$$

Dependent Variable[b]	Trade			
Independent Variables[c]	(1)	(2)	(3)	(4)
Earnings aggressiveness[d]	−0.0423 (0.0000)			
Loss avoidance[e]		0.0026 (0.4440)		
Earnings smoothing[f]			−0.0375 (0.0000)	
Overall earnings opacity[g]				−0.0496 (0.0000)
Liberalization[h]	0.2581 (0.0000)	0.2212 (0.0000)	0.2298 (0.0000)	0.2230 (0.0000)
Insider trading enforcement[i]	0.4535 (0.0000)	0.5010 (0.0000)	0.4813 (0.0000)	0.4855 (0.0000)
GDP growth[j]	−0.0030 (0.3613)	−0.0114 (0.0005)	−0.0051 (0.1197)	−0.0013 (0.7147)

---

[a] The numbers below are coefficient estimates from panel regressions, and are corrected for country-fixed effects, country-specific heteroscedasticity and country-specific autocorrelation. p-values are in parentheses.

[b] The dependent variable "trade" is defined as the natural logarithm of the ratio of volume of dollar trade per month to dollar market capitalization at the end of the month. These data were obtained from Datastream for the main stock market of each country. The sample periods for which these data were available are given in the Appendix.

[c] The first four independent variables are the earnings opacity variables, whereas the other independent variables are the control variables. The earnings opacity variables are rank variables. A higher rank implies more earnings opacity. The data to construct the earnings opacity variables come from Worldscope.

[d] We scale accruals by lagged total assets for each firm, determine its median in the cross-section of firms per country per year, and then rank these medians across years and across countries. This rank is the "earnings aggressiveness" time-series variable per country.

[e] We define firms with small positive earnings (small negative earnings) as firms with net income scaled by lagged total assets between 0 and 1 percent (between 0 and −1 percent). We subtract the number of firms with small negative earnings from the number of firms with small positive earnings per country per year, divide this difference by the sum of the two, and then rank this ratio across years and across countries. This rank is the "loss avoidance" time-series variable per country.

[f] We find the correlation between the change in accruals and the change in operating cash flows (both scaled by lagged total assets) in the cross-section of firms per country per year, and then rank these correlations across years and across countries. This rank is the "earnings smoothing" time-series variable per country.

[g] The "overall earnings opacity" time-series variable per country is the average of the "earnings aggressiveness" time-series variable per country, the "loss avoidance" time-series variable per country and the "earnings smoothing" time-series variable per country.

[h] The control variable "Liberalization" is an indicator variable. It changes from 0 to 1 in the month after the official liberalization date that was obtained from Bekaert and Harvey (2000). These liberalization dates are given in Table 1.

[i] The control variable "insider trading enforcement" is an indicator variable. It changes from 0 to 1 in the year after the first enforcement of insider trading laws. This date was obtained from Bhattacharya and Daouk (2002). These insider trading enforcement dates are given in Table 1.

[j] The control variable "GDP growth" is the growth rate of gross domestic product of a country every year. These data come from the World Bank, http://www.worldbank.org/research/growth/GDNdata.htm.

equations in the VAR is estimated *jointly*. This means that the effect of the independent variables on each endogenous variable takes into account the endogenous nature of the other endogenous variables.[21]

Formally, the system of equations to estimate the effect on the cost of equity is:

$$\begin{aligned} \textit{Cost of Equity}_{i,t} = {} & \beta_{10} + \beta_{11}\,\textit{Dimension of Earnings Opacity}_{i,t-1} \\ & + \beta_{12}\,\textit{Liberalization}_{i,t} + \beta_{13}\,\textit{Insider Trading Enforcement}_{i,t} \\ & + \beta_{14}\,\textit{GDP Growth}_{i,t} + \beta_{15}\,\textit{Market Capitalization/GDP}_{i,t-1} \\ & + u_{1i,t} \end{aligned}$$

and

$$\begin{aligned} \textit{Dimension of Earnings Opacity}_{i,t} = {} & \beta_{20} + \beta_{21}\,\textit{Cost of Equity}_{i,t-1} \\ & + \beta_{22}\,\textit{Liberalization}_{i,t} \\ & + \beta_{23}\,\textit{Insider Trading Enforcement}_{i,t} \\ & + \beta_{24}\,\textit{GDP Growth}_{i,t} \\ & + \beta_{25}\,\textit{Market Capitalization/GDP}_{i,t-1} \\ & + u_{2i,t} \end{aligned} \tag{5}$$

and the system of equations to estimate the effect on trade is:

$$\begin{aligned} \textit{Trade}_{i,t} = {} & \beta_{10} + \beta_{11}\,\textit{Dimension of Earnings Opacity}_{i,t-1} + \beta_{12}\,\textit{Liberalization}_{i,t} \\ & + \beta_{13}\,\textit{Insider Trading Enforcement}_{i,t} + \beta_{14}\,\textit{GDP Growth}_{i,t} \\ & + \beta_{15}\,\textit{Market Capitalization/GDP}_{i,t-1} + u_{1i,t} \end{aligned}$$

and

$$\begin{aligned} \textit{Dimension of Earnings Opacity}_{i,t} = {} & \beta_{20} + \beta_{21}\,\textit{Trade}_{i,t-1} + \beta_{22}\,\textit{Liberalization}_{i,t} \\ & + \beta_{23}\,\textit{Insider Trading Enforcement}_{i,t} \\ & + \beta_{24}\,\textit{GDP Growth}_{i,t} \\ & + \beta_{25}\,\textit{Market Capitalization/GDP}_{i,t-1} \\ & + u_{2i,t}. \end{aligned} \tag{6}$$

The system of equations is estimated jointly using Seemingly Unrelated Regressions (SUR). SUR computes estimates using the technique of joint GLS (Generalized Least Squares). The two error terms $u_{1i,t}$ and $u_{2i,t}$ are allowed to be correlated (see Enders [1996] for further details). The estimation allows for country-fixed effects, for country-specific heteroscedasticity, and for country-specific autocorrelation.

---

[21] A VAR is like a simultaneous equations model except that, instead of contemporaneous simultaneity, a lead-lag relationship is estimated. This lead-lag relationship is consistent with our empirical specification using lagged earnings opacity measures. The estimation procedure is akin to a 3SLS. The first two stages are similar to the 2SLS. The third stage takes into account that the error terms in the two equations are not independent. This third stage corrects for country-specific heteroscedasticity, and for country-specific autocorrelation.

We find that endogeneity does exist. Overall earnings opacity as well as all dimensions of earnings opacity, except earnings smoothing, are positively affected by the cost of equity using the dividend yield method. Earnings aggressiveness and earnings smoothing are affected by trade, but loss avoidance as well as overall earnings opacity are not affected by trade. However, though endogeneity exists and we explicitly account for it, our previous inferences about the effect of earnings opacity variables on cost of equity (using the dividend yield method) or trade are not affected.[22]

VI. CONCLUSION

This paper explores the link between earning opacity and cost of equity and trading volume in a broad cross-section of countries. We attempt to measure earnings opacity from the financial statements of firms. We use distributional properties of reported earnings to estimate for each country, for each year, three dimensions of earnings opacity: earnings aggressiveness, loss avoidance, and earnings smoothing. We combine these three dimensions to obtain an overall earnings opacity time-series measure per country. While not all results are consistent for our individual earnings opacity measures, we document in our panel data tests that, after controlling for other influences, an increase in overall earnings opacity in a country is linked to an increase in the cost of equity and a decrease in trading in the stock market of that country.

Our analyses have important limitations that should be kept in mind when interpreting our results. First, it is possible that earnings opacity, the cost of equity capital, and trading volume are all impacted by some unknown third variable, resulting in a spurious association between earnings opacity and our equity market measures. While we have attempted to control for the factors suggested by the past literature, theoretical and empirical limitations prevent us from knowing whether all of the important influences have been controlled. While our research design mitigates concerns over the endogeneity of earnings opacity, we are unable to ensure that endogeneity does not impact our analyses. Second, we undoubtedly measure earnings opacity with error, and this measurement error could impact our analyses. Finally, our cross-sectional tests relating earnings opacity to various factors that might influence financial reporting quality are cross-sectional tests rather than panel data tests. As such, these tests lack power due to the lack of data, and are particularly susceptible to bias caused by correlated omitted variables.

Keeping these limitations in mind, there are important implications of our analyses for investors, securities regulators, and academics. Our cross-sectional analysis documents associations between the proportion of auditors in the population and disclosure level and earnings opacity, suggesting that increased enforcement of accounting standards through auditing, and increased disclosure may improve earnings transparency. We document economically and statistically important relations between our measures of the distributional properties of reported earnings and the cost of equity capital and trade. These findings are consistent with the joint hypotheses that our measures are associated with opaque earnings, and that investors perceive risk associated with opaque earnings. So our result is consistent with the widely held belief that recent sharp declines in U.S. equity prices are in response to widely publicized accounting "scandals" in the United States, scandals that have heightened investors' concerns over earnings opacity, prompting investors to demand greater premiums.

---

[22] Because of the two-stage nature of the estimation of the cost of equity using the international asset pricing model method, it was not possible for us to run a VAR for this method.

Our analyses clearly suggest that further research into the impact of informational risk in general, and earnings opacity specifically, on equity markets is warranted. One avenue of future research could develop techniques to assess earnings opacity at the individual firm level, and then test for links between earnings opacity and equity market variables at the firm level rather than at a country level. A second avenue for research would be to determine which institutional factors impact which dimension of earnings opacity. A third avenue for research, and in our view the most useful avenue, is to develop more refined earnings opacity measures from the distribution of reported earnings.

APPENDIX
Description of Data Used

(1) Country	(2) Financial Statement Data (Annual) (Sample Period)	(3) Number of Firm-Years	(4) Indices of Stock Markets (Monthly) (Sample Period)	(5) Market Capitalization of Main Exchange (Monthly) (Sample Period)	(6) Dollar Volume in Main Exchange (Monthly) (Sample Period)	(7) Dividend Yield (Monthly) (Sample Period)	(8) GDP of Country (Quarterly or Annual) (Sample Period)	(9) Exports of Country (Monthly) (Sample Period)	(10) Imports of Country (Monthly) (Sample Period)	(11) Exchange Rate (Monthly) (Sample Period)
Australia	86Y–98Y	888	01/86–12/98	01/86–12/98	01/86–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
Austria	87Y–98Y	472	01/86–12/98	01/86–12/98	08/86–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
Belgium	86Y–98Y	567	01/86–12/98	01/86–12/98	01/86–12/98	01/86–12/98	86Y–98Y	01/93–12/98	01/93–12/98	01/86–12/98
Brazil	91Y–98Y	550	01/86–12/98	07/94–12/98	NA	NA	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Canada	86Y–98Y	1,997	01/86–12/98	01/86–12/98	01/86–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
Chile	94Y–98Y	147	01/86–12/98	07/89–12/98	07/89–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Denmark	88Y–98Y	953	01/86–12/98	01/86–12/98	04/88–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Finland	86Y–98Y	704	12/87–12/98	03/88–12/98	NA	03/88–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
France	86Y–98Y	3,834	01/86–12/98	01/86–12/98	06/88–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
Germany	86Y–98Y	3,847	01/86–12/98	01/86–12/98	06/88–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
Greece	90Y–98Y	491	01/86–12/98	01/88–12/98	01/88–12/98	01/90–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Hong Kong	87Y–98Y	925	01/86–12/98	01/86–12/98	06/88–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
India	92Y–98Y	1,342	01/86–12/98	01/90–12/98	01/95–12/98	01/90–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Indonesia	92Y–98Y	493	12/89–12/98	04/90–12/98	04/90–12/97	04/90–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Ireland	86Y–98Y	445	12/87–12/98	01/86–12/98	NA	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Italy	86Y–98Y	1,146	01/86–12/98	01/86–12/98	07/86–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Japan	86Y–98Y	8,762	01/86–12/98	01/86–12/98	01/90–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
Malaysia	86Y–98Y	1,233	01/86–12/98	01/86–12/98	01/86–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Mexico	90Y–98Y	361	01/86–12/98	01/88–12/98	01/88–12/98	05/89–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
The Netherlands	86Y–98Y	1,367	01/86–12/98	01/86–12/98	02/86–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Norway	88Y–98Y	502	01/86–12/98	01/86–12/98	01/86–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
Pakistan	92Y–98Y	361	01/86–12/98	NA	NA	NA	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Portugal	93Y–98Y	165	01/86–12/98	01/90–12/98	01/90–12/98	01/90–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Singapore	86Y–98Y	566	01/86–12/98	01/86–12/98	01/86–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98

(continued on next page)

APPENDIX (continued)

South Africa	86Y–98Y	889	12/92–12/98	01/86–12/98	01/90–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
South Korea	89Y–98Y	867	01/86–12/98	09/87–12/98	09/87–12/98	09/87–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
Spain	88Y–98Y	483	01/86–12/98	03/87–12/98	02/90–12/98	03/87–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Sweden	86Y–98Y	1,004	01/86–12/98	01/86–12/98	01/86–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Switzerland	86Y–98Y	1,261	01/86–12/98	01/86–12/98	01/89–12/98	01/86–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Taiwan	93Y–98Y	577	01/86–12/98	09/87–12/98	04/91–12/98	05/88–12/98	85Q4–98Y	01/88–12/98	01/88–12/98	12/93–12/98
Thailand	92Y–98Y	765	01/86–12/98	01/87–12/98	01/87–12/98	01/87–12/98	86Y–98Y	01/86–12/98	01/86–12/98	01/86–12/98
Turkey	93Y–98Y	188	12/86–12/98	01/88–12/98	01/88–12/98	06/89–12/98	87Q1–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
United Kingdom	86Y–98Y	8,974	01/86–12/98	01/86–12/98	10/86–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
United States	86Y–98Y	11,527	01/86–12/98	01/86–12/98	01/86–12/98	01/86–12/98	85Q4–98Q4	01/86–12/98	01/86–12/98	01/86–12/98
All Countries		58,653								

(1) Annual financial statement data for firms in 20 developed markets and 14 emerging markets were obtained from Worldscope. These countries are listed in Column (1). The sample period per country is given in Column (2). The number of firm-years is given in Column (3).

(2) Data on monthly stock market indices for the 20 developed markets were obtained from Morgan Stanley Capital Market International (MSCI). Data on monthly stock market indices for the 14 emerging markets were obtained from the International Financial Corporation (IFC). The sample periods are given in Column (4).

(3) Data on monthly market capitalization, dollar volume, and monthly dividend yields were obtained from Datastream. The sample periods are given in Columns (5), (6), and (7).

(4) Data on quarterly/annual GDP, monthly exports, monthly imports, and monthly foreign exchange rates were from the International Financial Statistics of the International Monetary Fund. The statistics for Taiwan come from Datastream. The sample periods are given in Columns (8), (9), (10), and (11).

REFERENCES

Alford, A., J. Jones, R. Leftwich, and M. Zmijewski. 1993. Relative informativeness of accounting disclosures in different countries. *Journal of Accounting Research* 31 (Supplement): 183–233.

Ali, A., and L-S. Hwang. 2000. Country-specific factors related to financial reporting and the value relevance of accounting data. *Journal of Accounting Research* 38: 1–21.

Amihud, Y., and H. Mendelson. 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics* 15: 223–249.

Ashbaugh, H., and M. Pincus. 2001. Domestic accounting standards, international accounting standards, and the predictability of earnings. *Journal of Accounting Research* 39: 417–434.

Ball, R., S. P. Kothari, and A. Robin. 2000. The effect of international institutional factors on properties of accounting earnings. *Journal of Accounting and Economics* 29: 1–51.

———, A. Robin, and J. S. Wu. 2000. Incentives versus standards: Properties of accounting income in four East Asian countries, and implications for acceptance of IAS. Working paper # FR00-04, University of Rochester.

Barth, M. E., J. A. Elliott, and M. W. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37 (Autumn): 387–413.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24: 3–37.

Bekaert, G., and C. Harvey. 1995. Time varying world market integration. *Journal of Finance* 50: 403–444.

———, and ———. 1997. Emerging equity market volatility. *Journal of Financial Economics* 43: 29–77.

———, and ———. 2000. Foreign speculators and emerging equity markets. *Journal of Finance* 55: 565–613.

Bhattacharya, U., and M. Spiegel. 1991. Insiders, outsiders, and market breakdowns. *Review of Financial Studies* 4: 255–282.

———, and H. Daouk. 2002. The world price of insider trading. *Journal of Finance* 57: 75–108.

Bollerslev, T., R. Engle, and J. Wooldrige. 1988. A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96: 116–131.

Brennan, M., and A. Subrahmanyam. 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics* 41: 441–464.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24: 99–126.

Bushman, R., and A. Smith, 2001. Financial accounting information and corporate governance. *Journal of Accounting and Economics* 32: 237–333.

Center for International Financial Analysis and Research (CIFAR). 1995. *International Accounting and Auditing Trends*. 4th Edition. Edited by V. B. Bavishi. Princeton, NJ: CIFAR.

Chang, J. J., T. Khanna, and K. G. Palepu. 2000. Analyst activity around the world. Working paper, Philadelphia, PA.

Choi, F., C. Frost, and G. K. Meek. 1999. *International Accounting*. 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70: 193–226.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72: 1–33.

Demirguc-Kunt, A., and V. Maksimovic. 1998. Law, finance, and firm growth. *Journal of Finance* 53: 2107–2137.

Dumas, B., and B. Solnik. 1995. The world price of foreign exchange risk. *Journal of Finance* 50: 445–479.

Enders, W., 1996. *RATS Handbook for Econometric Time Series*. New York, NY: John Wiley & Sons.

Engle, R., D. Lilien, and R. Robins. 1987. Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica* 55: 391–407.

Ferson, W., and C. Harvey. 1993. The risk and predictability of international equity returns. *Review of Financial Studies* 6: 527–566.

Financial Accounting Standards Board (FASB). 1978. *Objectives of Financial Reporting by Business Enterprises*. Statement of Financial Accounting Concepts No 1. Norwalk, CT: FASB.

Fudenberg, D., and J. Tirole. 1995. A theory of income and dividend smoothing based on incumbency rents. *Journal of Political Economy* 103: 75–93.

Glosten, L., and P. Milgrom. 1985. Bid, ask and transaction prices in a specialist model with heterogeneously informed traders. *Journal of Financial Economics* 14: 71–100.

Harris, T. S., M. Lang, and H. P. Möller. 1994. The value relevance of German accounting measures: An empirical analysis. *Journal of Accounting Research* 32 (Autumn): 187–209.

Harvey, C. 1991. The world price of covariance risk. *Journal of Finance* 46: 111–157.

———. 1995. Predictable risk and returns in emerging markets. *Review of Financial Studies* 8: 773–816.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7: 85–107.

Henry, P. 2000. Stock market liberalization, economic reform, and emerging market equity prices. *Journal of Finance* 55: 529–564.

Himmelberg, C., R. G. Hubbard, and D. Palia. 1999. Understanding the determinants of managerial ownership and the link between ownership and performance. *Journal of Financial Economics* 53: 353–384.

International Accounting Standards Committee (IASC). 1997. *Insight*. London, U.K.: IASC.

Jacoby, G., D. Fowler, and A. Gottesman. 2000. The capital asset pricing model and the liquidity effect: A theoretical approach. *Journal of Financial Markets* 3: 69–81.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (Autumn): 193–228.

Joos, P., and M. Lang. 1994. The effects of accounting diversity: Evidence from the European Union. *Journal of Accounting Research* 32 (Supplement): 141–168.

Kyle, A. 1985. Continuous auctions and insider trading. *Econometrica* 53: 1315–1335.

Land, J., and M. Lang. 2002. Empirical evidence on the evolution of international earnings. Working paper, The University of North Carolina at Chapel Hill, Chapel Hill, NC.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny. 1997. Legal determinants of external finance. *Journal of Finance* 52: 1131–1150.

———, ———, ———, and ———. 1998. Law and finance. *Journal of Political Economy* 106: 1113–1155.

Leuz, C., D. Nanda, and P. Wysocki. 2003. Earnings management and institutional factors: An international comparison. *Journal of Financial Economics* (forthcoming).

Levine, R. 1997. Financial development and economic growth: Views and agenda. *Journal of Economic Literature* 35: 688–726.

PricewaterhouseCoopers. 2001. *The Opacity Index*. New York, NY: PricewaterhouseCoopers.

Rangan, S. 1998. Earnings management and the performance of seasoned equity offerings. *Journal of Financial Economics* 50: 101–122.

Saudagaran, S, and J. Diga. 1997. Financial reporting in emerging capital markets: Characteristics and policy issues. *Accounting Horizons* 11: 41–64.

Shivakumar, L. 2000. Do firms mislead investors by overstating earnings before seasoned equity offerings? *Journal of Accounting and Economics* 29: 339–371.

Sims, C. 1980. Macroeconomics and reality. *Econometrica* 48: 1–49.

Stulz, R. 1999a. Globalization of equity markets and the cost of capital. Working paper, National Bureau of Economic Research.

———. 1999b. Globalization, corporate finance, and the cost of capital. *Journal of Applied Corporate Finance* (Fall): 8–25.

Teoh, S. H., I. Welch, and T. J. Wong. 1998. Earnings management and the underperformance of seasoned equity offerings. *Journal of Financial Economics* 50: 63–99.

Trueman, B., and S. Titman. 1988. An explanation for accounting income smoothing. *Journal of Accounting Research* 26 (Supplement): 127–139.

World Factbook. 2001. Central Intelligence Agency, VA. Available at http://www.odci.gov/cia/publications/factbook.

Zarowin, P. 2002. Does income smoothing make stock prices more informative? Working paper, New York University, New York, NY.

Zhou, X. 2001. Understanding the determinants of managerial ownership and the link between ownership and performance: comment. *Journal of Financial Economics* 62: 559–571.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 679–706

Sophistication-Related Differences in Investors' Models of the Relative Accuracy of Analysts' Forecast Revisions

Sarah E. Bonner
University of Southern California

Beverly R. Walther
Northwestern University

Susan M. Young
Emory University

ABSTRACT: The accuracy of sell-side analysts' forecast revisions is related to a number of factors, including characteristics of the analyst and the age of the forecast. In this study we examine whether there are differences in how sophisticated and unsophisticated investors use these factors to predict the relative accuracy of forecast revisions. We adapt the lens model methodological approach from the judgment and decision-making literature to investigate these differences in an archival setting. Our results suggest that sophisticated investors have greater knowledge overall about the relation of the factors to forecast accuracy. Further, our evidence is consistent with sophisticated investors relying more on the specific factors that provide the most benefits (relative to their costs) for predicting relative forecast accuracy.

Keywords: *analyst forecast revisions; market reaction; investor sophistication; lens model.*

Data Availability: *Data are available from public sources.*

I. INTRODUCTION

Prior research has shown that the accuracy of sell-side analysts' forecast revisions is systematically associated with characteristics of the analyst and the age of the forecast (e.g., O'Brien 1988; Stickel 1992; Clement 1999; Jacob et al. 1999). In this study,

---

We thank Linda Bamber, Jan Barton, Michael Clement, Joel Demski, Artur Hugon, Hai Lu, Mark Nelson, Grace Pownall, Leonard Soffer, Geoff Sprinkle, Kristy Towry, Sydney S. Y. Young, two anonymous reviewers, participants at the Behavioral-Financial Research Conference at Emory University, and workshop participants at Michigan State University, Southern Methodist University, University of California, Irvine, University of Illinois at Chicago, and Washington University for helpful comments and discussions. We appreciate the financial support, respectively, of BDO Seidman, LLP and the Marshall School of Business, the Kellogg School of Management, and the Goizueta Business School. We are grateful for the research assistance of Artur Hugon and Wai-Ming Loke. Earnings forecasts are from Zacks Investment Research. Errors or omissions are our responsibility.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted August 2001
Accepted January 2003

we examine the extent to which investors consider these factors when reacting to forecast revisions. Specifically, we investigate whether there are differences between sophisticated and unsophisticated investors in how they use these factors to predict the relative accuracy of analysts' forecast revisions. We first examine whether sophisticated investors have greater knowledge overall about the factors associated with forecast accuracy. We then examine whether sophisticated investors are better able to identify factors that provide large benefits (relative to their costs) for predicting forecast accuracy, and therefore focus more on these factors.[1]

We address these questions by examining the quality of the models investors implicitly use to predict relative forecast accuracy. We infer sophisticated and unsophisticated investors' prediction models from their stock price reactions to analyst forecast revisions. We then compare these prediction models to a statistical model, which specifies relative forecast accuracy as a function of the factors shown by previous research to be associated with forecast accuracy. We base our inferences regarding differences between sophisticated and unsophisticated investors' prediction models on the "matching index" from Brunswik's lens model of human judgment. The matching index represents the correlation between the fitted values from investors' (implied) models of relative forecast accuracy and the fitted values from the statistical model of relative forecast accuracy. Our measure of investor sophistication is based on five firm characteristics: analyst following, percentage of institutional ownership, number of institutions holding shares, number of shares held by institutions, and dollar value of shares traded. We use a factor analysis of these characteristics to obtain a single measure of investor sophistication for each firm-year observation, and then compute matching indexes separately for sophisticated and unsophisticated investors.

Our first hypothesis is that sophisticated investors have greater knowledge overall about how to use the information in the set of factors that can be used to predict analysts' forecast accuracy. To test this hypothesis, we compare the matching indexes for sophisticated and unsophisticated investors. As predicted, the matching index for sophisticated investors is statistically higher than the matching index for unsophisticated investors.

Our second hypothesis is that sophisticated investors are more knowledgeable about the benefits of individual factors for predicting analysts' relative forecast accuracy. To test this hypothesis, we compare the matching indexes from a model that includes only the two most beneficial factors for predicting forecast accuracy (the age of the forecast and the analyst's prior accuracy) to those from more comprehensive models that also include less beneficial factors. If sophisticated investors ignore or underweight the information in these less beneficial factors more than unsophisticated investors, the decrease in their matching index will be statistically greater than the decrease in unsophisticated investors' matching index when these factors are added to the model. The results are consistent with our prediction, suggesting that sophisticated investors exhibit more "adaptive decision-making." Taken together, the findings suggest that sophisticated investors not only have more knowledge overall about the set of factors related to analysts' forecast accuracy, but they also have greater knowledge of the individual factors that are most beneficial to use.

Our findings contribute to the literature in the following ways. First, we provide information about how different types of investors use analysts' forecasts when predicting the earnings of a firm. Sell-side analysts' earnings forecasts are an important source of information for investors making predictions about the future earnings and profitability of a firm and investment decisions based on these predictions (SRI International 1987; Williams et al. 1996; Hodge 2001). While recent research has examined investors' reactions to analysts' forecast revisions conditional on factors that are related to forecast accuracy (Stickel 1992; Mikhail et al. 1997; Park and Stice 2000; Clement and Tse 2003; Gleason and Lee 2003), no previous study has examined sophistication-related differences in these reactions. Investigating sophistication-related differences in investors' reactions, specifically their implied prediction models for the accuracy of sell-side analysts' earnings forecasts, can provide insights into whether and how unsophisticated investors' decisions can be improved. Additionally, our factor analysis of various sophistication proxies may provide guidance to researchers interested in examining other sophistication-related differences in investor behavior.

---

[1] Although we use the terminology "sophisticated" and "unsophisticated" investors, we recognize that our "unsophisticated" investors are not necessarily untrained, but are simply less sophisticated than those we define as "sophisticated."

Our study further contributes to the literature in that we illustrate how a methodological approach from behavioral research (the lens model) can be used to study investor behavior in an archival setting. Because this approach focuses on the extent to which the *output* from investors' implied prediction models matches that from the statistical model, it is appropriate in any setting in which investors are making predictions using factors that are correlated. The lens model accounts for the fact that investors can exploit the correlation among factors and, consequently, make good predictions even if their models contain inaccurate individual factor weightings. However, as with any technique that infers investors' prediction models from market reactions, our approach requires several assumptions. While additional analyses indicate that the assumptions likely hold in our setting, we are unable to rule out all alternative explanations for our findings.

This paper is organized as follows. The next section reviews the literature, followed by a description of our methodological approach in Section III. Section IV develops our hypotheses. Section V discusses the sample, variables, and models we use. Section VI presents the results of our analyses. Section VII contains robustness tests, including analyses to address the validity of the assumptions underlying our empirical tests. Section VIII concludes.

II. LITERATURE REVIEW

Sell-side analysts' earnings forecasts and forecast revisions are an important source of information for investors making predictions about the future earnings and profitability of a firm and investment decisions based on these predictions (SRI International 1987; Williams et al. 1996; Hodge 2001). Prior research has documented that there is a significant stock market reaction associated with analyst forecast revisions (see, for example, Gonedes et al. 1976; Givoly and Lakonishok 1979; Brown et al. 1985). Overall, this literature indicates that sell-side analysts' forecast revisions provide information to investors.

More recent research has controlled for the sign and magnitude of the forecast revision, and documented that investors' reaction to these revisions varies with factors that are associated with the accuracy of the forecasts. Stickel (1992) finds that there is a stronger price reaction to upward, but not downward, forecast revisions issued by analysts named to the *Institutional Investor* All-American team (see also Gleason and Lee 2003). Park and Stice's (2000) and Chen et al.'s (2001) results indicate that investors react more strongly to forecast revisions issued by analysts with high prior forecasting accuracy. Finally, in an experimental study, Maines (1996) finds that investors consider the historical accuracy of analysts' forecasts when combining them to derive a consensus forecast. These findings suggest that investors appear to be attempting to predict the accuracy of analyst forecasts when reacting to them.[2]

In contrast to prior studies that examined investors' reaction to analyst forecast revisions conditional on only one factor associated with forecast accuracy, Clement and Tse (2003) examine investors' reaction conditional on multiple factors associated with forecast accuracy. They find that the association of investors' reaction with some factors is inconsistent with those factors' association with forecast accuracy. For example, their results indicate that investors react more strongly to forecasts issued earlier; however, earlier forecasts tend to be less accurate (e.g., O'Brien 1990). Consequently, their work suggests that investors do not properly incorporate all of the information in factors for forecast accuracy when reacting to forecast revisions.

Prior studies of investors' reaction to forecast revisions are somewhat limited in providing insights about the quality of investors' prediction of forecast accuracy for three reasons. First, some prior studies have documented that investors' reaction to forecast revisions is a function of factors associated with the accuracy of these revisions, but have not directly assessed the quality of investors' prediction model by examining, for example, whether investors' weightings on the factors correspond to the statistical association between the factors and forecast accuracy (e.g., Stickel 1992). Other studies have investigated the quality of investors' prediction model, but have used the Mishkin (1983) technique to examine the similarity of individual coefficients on factors in the model of investors' reaction to forecast revisions to the corresponding coefficients in the model of forecast accuracy (e.g., Clement and Tse 2003). By contrast, we use the lens model, described in detail in the next section, to assess the quality of investors' prediction model. This approach focuses on the similarity of *fitted values* from investors' prediction model of forecast accuracy to the *fitted values* obtained from a statistical model of forecast accuracy. The lens model implicitly allows for a setting in which investors exploit correlation among the factors by testing the accuracy of the *output* from investors' model rather than the accuracy of their weightings on individual factors.

Second, prior studies on the relation between investors' reaction to forecast revisions and the factors associated with forecast accuracy do not allow for heterogeneity in how investors incorporate the information in factors for forecast accuracy when predicting the accuracy of a revised forecast. However, recent studies have suggested that a firm's stock price reaction to information is associated with proxies for the sophistication of the firm's marginal investor. For example, Bartov et al. (2000) find that the magnitude of the post-earnings announcement drift decreases as the sophistication of the firm's investors (proxied by the percentage of stock held by institutions) increases (also see Hand 1990). We examine differences in the quality of investors' prediction models for forecast accuracy between subsamples formed on the basis of the sophistication of the marginal investor. Finally, there are multiple facets of the "quality" of prediction models. In addition to examining the quality of models in terms of the accuracy of the model's overall predictions for forecast accuracy, we examine whether sophisticated investors better consider the costs and benefits of using individual factors in constructing their models than do unsophisticated investors. In the next two sections, we describe the methodology we use to investigate these issues and our hypotheses concerning these differences in model quality.

---

[2] Additionally, Brown and Mohammed (2001) find that investors can develop a profitable trading strategy based on an understanding of the accuracy of analysts' forecasts.

III. METHODOLOGY

To investigate differences between sophisticated and unsophisticated investors as to the quality of their models for predicting analysts' forecast accuracy, we use Brunswik's lens model. Brunswik (and later Hammond) developed the lens model to examine how well individuals assess the relation between factors and outcomes in the environment in making predictions about future outcomes (see Hammond and Stewart [2001] for a review of Brunswik's work). In the traditional application of the lens model, individuals are given a set of factors and are asked to make predictions of outcomes (e.g., analyst forecast accuracy) in a laboratory setting. Using these data, the researcher would calculate a statistical model and each individual decision-maker's model of analyst forecast accuracy, both based on the given set of factors. The correlation between the fitted values from the statistical model and the fitted values from an individual decision-maker's model is called the matching index (see Ashton [1982] for an excellent discussion of the lens model). The matching index measures the quality of the linear component of the decision-maker's model, or, put more simply, the accuracy of the factor signs and weights *only* to the extent they affect the accuracy of the model's output (e.g., predictions of analyst forecast accuracy).[3]

Adaptation of the Lens Model

The lens model was developed to study individuals' predictions. By contrast, in our setting, we have archival market-level data, requiring us to infer investors' predictions from market prices. Many assumptions must hold in order for our application of this technique to yield valid inferences. For example, we must assume that all investors have equal access both to analyst forecast revisions and to the factors that can be used to predict their accuracy. In Section VII, we outline the necessary assumptions, provide empirical evidence on the appropriateness of these assumptions where possible, and assess the sensitivity of our results to these assumptions.

To illustrate how we adapt the lens model to our setting, assume for ease of exposition that there is only one variable that is informative for predicting the accuracy of the analyst's forecast revision, the analyst's prior accuracy relative to all other analysts who follow the firm (*PRMAPE*). Therefore, the equation for the relative accuracy of the analyst's forecast revision is:

$$PRMAPE_{i,j,q} = \alpha_0 + \alpha_1 PRMAPE_{i,j,q-1} + \varepsilon_{i,j,q} \tag{1}$$

where:

$PRMAPE_{i,j,q}$ = the absolute value of analyst *i*'s forecast error for firm *j* and quarter *q* minus the mean absolute forecast error for firm *j* and quarter *q*, deflated by the mean absolute forecast error for firm *j* and quarter *q*; and

$\varepsilon_{i,j,q}$ = error term with assumed zero mean and constant variance.

The fitted values from Equation (1) are the statistical model's predictions of the relative accuracy of the analyst's forecast revision ($PR\hat{M}APE^{STAT}_{i,j,t}$).

---

[3] Many studies calculate "achievement" rather than or in addition to the matching index. Achievement is the correlation of an individual's predictions and the actual outcomes, and is influenced by noise in the environment (how highly associated with the outcome the set of factors is), matching of factor signs and weights, and noise in the individual's use of his prediction policy. We focus solely on the matching component of overall achievement in order to examine directly the quality of investors' models.

Further assume that there is only one type of investor, and that investors consider the predicted relative accuracy of an analyst's earnings forecast revision when reacting to its release:[4]

$$CAR_{i,j,t} = \beta_0 + \beta_1 FCREV_{i,j,t} \cdot PRMAPE^{PRED}_{i,j,t} + \xi_{i,j,t} \tag{2}$$

where:

$CAR_{i,j,t}$ = three-day ($t-1$, $t+1$) size-adjusted return surrounding the issue of the revised quarterly earnings forecast by analyst i for firm j at time t;

$FCREV_{i,j,t}$ = the price-deflated revision in analyst i's forecast for firm j issued at time t;

$PRMAPE^{PRED}_{i,j,t}$ = investors' predicted relative accuracy of analyst i's forecast for firm j issued at time t; and

$\xi_{i,j,t}$ = error term with assumed zero mean and constant variance.

In the lens model framework, the matching index is the correlation between the fitted values from the statistical model and the fitted values from an individual decision-maker's model. Loosely speaking, this is equivalent to calculating the correlation between $PRMAPE^{PRED}_{i,j,t}$ in Equation (2) with the fitted values, $PR\hat{M}APE^{STAT}_{i,j,t}$, from Equation (1).

To infer investors' prediction of the relative accuracy of the forecast revision, we substitute Equation (1) into Equation (2):

$$CAR_{i,j,t} = \beta_0 + \beta_1 FCREV_{i,j,t} \cdot (\alpha'_0 + \alpha'_1 PRMAPE_{i,j,q-1}) + \xi_{i,j,t} \tag{3}$$

where α'_0 and α'_1 equal investors' weightings used to predict the relative accuracy of the analyst's forecast revision. Therefore, to infer investors' prediction of the relative accuracy of the forecast revision ($PR\hat{M}APE^{INV}_{i,j,t}$), we replace the estimated coefficients in Equation (1) with these estimated α''s from Equation (3). The matching index can be expressed as:

$$Matching\ Index = corr\ (PR\hat{M}APE^{STAT}_{i,j,t},\ PR\hat{M}APE^{INV}_{i,j,t}). \tag{4}$$

The lens model has not been used previously in an empirical study of market-level prices, and differs from the Mishkin (1983) approach traditionally used in the literature.[5] We do not use the Mishkin approach because, as discussed in the next section, our hypotheses rely on the existence of different types of investors (sophisticated and unsophisticated) and predict that sophisticated investors' model is of higher quality than unsophisticated investors' model. One reason this result could occur is if sophisticated investors' factor weightings are closer to the weights from a statistical model of forecast accuracy than unsophisticated investors' weightings.[6] In the extreme, we could use the Mishkin (1983) approach to test whether sophisticated investors' α''s from Equation (3) equal the α's from Equation (1), and test separately whether unsophisticated investors' α''s from Equation (3) equal the α's from Equation (1). However, if we reject the null hypothesis that $\alpha = \alpha'$ for both investor types, then we have not necessarily rejected (or failed to reject) the null hypothesis that the quality of sophisticated investors' model of the relative accuracy of an analyst's forecast revision is equal to the quality of unsophisticated investors' model. The lens model allows us to compare the models of sophisticated and unsophisticated investors without this difficulty.

---

[4] Consistent with behavioral studies that use the lens model methodology, we assume investors use a linear decision model.

[5] The Mishkin (1983) approach tests whether the estimated coefficients from the statistical model (the α's from Equation (1)) equal the estimated coefficients from the investor model (e.g., the α''s from Equation (3)).

[6] Stated in the context of Equations (1) and (3), if sophisticated investors' α 's from Equation (3) are closer to the α''s from Equation (1) than are unsophisticated investors' α's from Equation (3) (i.e., $|\alpha'^{SOPH} - \alpha| < |\alpha'^{UNSOPH} - \alpha|$), then the fitted values derived from sophisticated investors' model would more closely match the fitted values derived from the statistical model.

Second, research has identified factors aside from the analyst's prior accuracy that are related to the relative accuracy of analysts' forecast revisions (e.g., Clement 1999; Jacob et al. 1999; Brown 2001). The environment in which analysts work is one that creates correlation among these factors. For example, firm-specific experience and general experience tend to be highly correlated; this occurs because analysts tend to cover firms for lengthy periods of time rather than periodically switching the firms they cover. Consequently, investors may weight correlated factors (Equation (3)) in a manner that differs from their weighting in the statistical model (Equation (1)) but still make predictions of the relative accuracy of analysts' forecast revisions that are similar to those that would be derived from the statistical model. For example, an investor who relies mostly on firm-specific experience when predicting forecast accuracy (and, thus, puts little weight on general experience) could make very similar predictions to those derived from a statistical model that relies mostly on general experience. As discussed in the previous section, the lens model methodology implicitly allows for the situation in which investors exploit the correlation among the factors in the environment by focusing on investors' overall predictions, rather than their weightings of individual factors. In essence, the lens model methodology accounts for differences in factor signs and weights between investors and the statistical model, but only to the extent they have an effect on the accuracy of investors' predictions. Further, to the extent we omit factors that are correlated with the factors we include, the fitted values from investors' models likely reflect the use of these correlated factors.

While the lens model and its matching index have many desirable properties for testing our hypotheses, as discussed above, many assumptions must hold in order for the application of a technique that seeks to infer investor behavior from market reaction data to yield valid inferences. In essence, there must be a link between individual investors' predictions of the accuracy of an analyst forecast based on observable factors and the market reaction to that forecast revision. Therefore, the conclusions we draw from our empirical results are subject to the assumptions necessary for our methodological approach, as outlined in Section VII.

IV. HYPOTHESES

We use the lens model to investigate two differences in the quality of sophisticated and unsophisticated investors' (implied) models for predicting the accuracy of analysts' forecast revisions. Our first hypothesis is that sophisticated investors have greater knowledge overall than unsophisticated investors about how to use the information in the set of factors that predict relative forecast accuracy, resulting in a higher matching index for sophisticated investors than for unsophisticated investors. Sophisticated investors may have greater knowledge because they have more experience with analyst forecasts through forecasting earnings on a full-time basis (Yunker and Krehbiel 1988; Potter 1992) and through following a larger number of stocks (Barber and Odean 2000); more experienced individuals have more knowledge about factor signs and weights (e.g., Slovic 1969; Einhorn 1974; Johnson et al. 1981; Johnson 1988; Bonner 1990). Further, sophisticated investors may be better able to learn the appropriate signs and weights for factors associated with forecast accuracy because they have superior abilities (Beaver 1998), and people with superior abilities learn better from experience (Ackerman et al. 1989; Bonner and Walker 1994). In addition, sophisticated investors may have access to resources that assist them in learning about analysts' forecasts (such as analyst rating services) that unsophisticated investors do not. Finally, sophisticated investors may have greater incentives to learn about the factors associated with analysts' forecast accuracy since their decisions based on analysts' forecasts likely have greater financial consequences, assuming they manage larger portfolios (Cready 1988; Hassel and Norman 1992; Brous and Kini 1994; El-Gazzar 1998). Therefore, our first hypothesis, stated in alternative form, is:

> **H1:** The matching index for sophisticated investors' model of relative forecast accuracy will be greater than the matching index for unsophisticated investors' model of relative forecast accuracy.

Our second hypothesis concerns differences in the degree to which sophisticated and unsophisticated investors' models reflect cost-benefit trade-offs related to the individual factors associated with forecast accuracy. Psychology research suggests people will employ "adaptive decision making" by choosing factor-weighting strategies based on their expected benefits and costs (Bettman et al. 1990; Payne et al. 1993). The key potential benefit from correctly incorporating information in a given factor is making superior predictions of the accuracy of forecasts; the costs include the cognitive effort of acquiring and processing the information needed to use the factor in a prediction model. The net benefit of factors investors use to predict analyst forecast accuracy can vary substantially. For example, the age of the forecast is very beneficial for predicting its accuracy (e.g., O'Brien 1988; Brown 2001), and it is not costly to obtain and use—the investor simply needs to know the date on which the forecast is issued and the time period to which it pertains. By contrast, the explanatory power of the number of firms followed by the analyst for forecast accuracy is not as great, and the costs of acquiring and using this information are greater—the investor must gather information on all firms the analyst follows, regardless of whether the investor follows all the firms himself.

We expect that sophisticated investors' model will better reflect the net benefit of using individual factors because they have greater knowledge of the net benefit of the factors that predict forecast accuracy. The psychology literature suggests that the sophistication (experience) of the decision maker may create differences in factor-weighting strategies because experience affects the accuracy of the benefit expectations. Payne et al. (1993) and Fennema and Kleinmuntz (1995) note that people must have extensive experience making and receiving feedback about their judgments in order to learn to anticipate the extent to which focusing on various factors results in better predictions. These studies also have shown few sophistication-related differences in expectations about the costs of using factors for predictions. In our setting, this finding implies that sophisticated and unsophisticated investors may be equally knowledgeable about the costs of acquiring and processing various factors, but that there will be sophistication-related differences in investors' knowledge of the expected benefits. These differences would lead to the investor model better distinguishing between high net benefit factors and low net benefit factors when the marginal investor is more likely to be sophisticated.

To illustrate how we derive our second hypothesis, consider an extreme case in which sophisticated investors only focus on the two most beneficial factors (e.g., forecast age and prior accuracy) to predict relative forecast accuracy, and ignore all other factors. When the statistical model includes only forecast age and prior accuracy, sophisticated investors' matching index should be very high. When another (less beneficial and more costly) factor (e.g., number of firms followed) is added to the statistical model, we predict that sophisticated investors' matching index will decrease because they ignore or underweight the information in the number of firms followed. Therefore, to examine whether sophisticated investors' model better reflects cost-benefit trade-offs related to the factors associated with forecast accuracy, we examine whether sophisticated investors' matching index decreases more than unsophisticated investors' matching index as factors with increasingly lower net benefit are added to the statistical model.[7] Our second hypothesis, stated in alternative form, is:

> **H2:** The matching index for sophisticated investors' model of relative forecast accuracy will decrease more as factors lower in net benefit are added to a statistical model that includes only high-benefit factors than will the matching index for unsophisticated investors' model of relative forecast accuracy.

The next section discusses the data we use to test our hypotheses.

V. DATA AND DESCRIPTIVE STATISTICS

Sample Selection

Our initial sample contains 208,240 revisions in quarterly earnings forecasts from the Zacks Investment Research database for the period 1981–1999. These observations represent forecast revisions issued closest in time to the quarterly earnings announcement date by individual analysts with sufficient data to calculate our dependent and independent variables (described below).[8] We eliminate 5,625 revisions that are less than five calendar days before the quarterly earnings announcement. We impose this minimum forecast horizon so that we can attribute the market reaction to the forecast revision and not the release of actual earnings. We also require that the revised forecast be issued after the beginning of the fiscal quarter to which it pertains to ensure that our results are not due to the inclusion of "stale" forecasts in our sample. This requirement eliminates 47,432 observations.[9] We control for time-period and firm effects in our empirical tests by mean-adjusting each variable; this procedure eliminates the need to include any firm characteristics or time-period variables in the regression model (see Clement 1999). Because we mean-adjust each variable, we further require a minimum of two analysts for each firm and quarter. This requirement eliminates 25,744 observations. We eliminate 22,863 observations with missing values for our proxies for investor sophistication (described below). Finally, to minimize the effect of outliers on our results, we eliminate 5,338 observations in the extreme 1 percent of the distribution of our two dependent variables, forecast accuracy and the market reaction to the forecast revision (see Clement 1999).[10] Our final sample includes 101,238 observations, representing 3,290 analysts and 1,757 firms from 1981 to 1999.

---

[7] We predict that the benefits and costs of incorporating the information in various factors when predicting analysts' forecast accuracy vary across *factors*. However, the benefits and costs of incorporating information from various factors also may vary between sophisticated and unsophisticated *investors*. It is reasonable to expect that sophisticated investors may have greater benefits from making good predictions about the accuracy of analysts' forecasts (Cready 1988; Hassel and Norman 1992; Brous and Kini 1994; El-Gazzar 1998) and/or lower costs of acquiring and processing factors to make predictions about forecast accuracy (Wilson 1975; Lev 1988). However, this would bias against finding results consistent with our hypothesis. If, for example, the net benefit of using the number of firms followed is greater for sophisticated investors than for unsophisticated investors, then we would expect a *smaller* decrease (or larger increase) in sophisticated investors' matching index compared to unsophisticated investors' matching index when this factor is added to the statistical model.

[8] We eliminate forecast revisions that are associated with 345 analyst codes on Zacks that correspond either to an unidentified individual (e.g., Value Line) or a brokerage house. Our results are unchanged if we retain all quarterly forecast revisions issued by individual analysts in our analysis instead of only the one issued closest in time to the quarterly earnings announcement date.

[9] Our results are unchanged if we retain forecasts issued prior to the beginning of the fiscal quarter in our analyses.

Variable Definitions and Descriptive Statistics

Table 1 provides descriptive statistics for our variables. The mean (median) absolute percentage forecast error, defined as the absolute value of actual quarterly earnings per share minus the quarterly earnings forecast, deflated by price ten trading days before the release of the forecast, is 0.63 percent (0.17 percent). Consistent with prior findings that analysts revise their forecasts downward throughout the period (Richardson et al. 2001), the forecast revisions tend to be negative. The mean forecast revision (*FCREV*), defined as the revised forecast minus the prior forecast, deflated by price ten trading days before the release of the revised forecast, is –0.0016. The median forecast revision, however, is 0.0000. Consistent with the negative mean forecast revision, the market reaction at the forecast

TABLE 1
Descriptive Statistics for Quarterly Forecast Revisions (n = 101,238)

Variable[a]	Mean	Median	First Quartile	Third Quartile	Std. Dev.
MAPE	0.0063	0.0017	0.0006	0.0058	0.0130
FCREV	−0.0016	0.0000	−0.0021	0.0000	0.0059
CAR	−0.0021	−0.0017	−0.0234	0.0209	0.0497
FCAGE	49.7878	44.0000	22.0000	78.0000	31.0522
FIRMEXP	13.8311	10.0000	6.0000	19.0000	11.1747
GENEXP	327.3060	244.0000	113.0000	452.0000	311.0560
TURNOVER	0.1235	0.0000	0.0000	0.0000	0.3290
FCFREQ	4.6819	4.0000	3.0000	6.0000	2.5195
NOFIRM	18.0610	16.0000	12.0000	21.0000	10.3890
NOIND	5.2896	5.0000	3.0000	7.0000	3.2588
BROKSIZE	41.1839	35.0000	18.0000	58.0000	28.5140
IIAWARD	0.2161	0.0000	0.0000	0.0000	0.4116

[a] Variable definitions: The Mean Absolute Percentage Error (*MAPE*) is the absolute value of actual quarterly earnings per share minus the quarterly earnings forecast, deflated by price ten trading days before the release of the forecast. *FCREV* is the revised forecast minus the prior forecast, deflated by price ten trading days before the release of the revised forecast. *CAR* is the three-day size-adjusted return centered on the forecast release date. *FCAGE* is the number of calendar days between the firm's earnings announcement date and the forecast release date. *FIRMEXP* is the number of prior quarters for which the analyst has issued an earnings forecast for the firm. *GENEXP* is the number of prior firm quarters the analyst has issued a quarterly forecast for any firm on the Zacks database. *TURNOVER* is an indicator variable that equals 1 if the analyst changed brokerage houses during the year, and 0 otherwise. *FCFREQ* is the number of quarterly earnings forecast issued by the analyst for the firm and quarter. *NOFIRM* is the number of firms for which the analyst issues quarterly earnings announcement forecasts for the quarter. *NOIND* is the number of two-digit SIC codes containing firms for which the analyst issues quarterly earnings announcements. *BROKSIZE* is the number of analysts issuing quarterly earnings forecasts at the brokerage house at which the analyst is employed. *IIAWARD* is an indicator variable that equals 1 if the analyst was named to the *Institutional Investor* All-American team in the previous year.

---

[10] Our results are not sensitive to using alternative outlier detection techniques, such as eliminating observations on the basis of studentized residuals or Cook's D in the regression estimations.

revision date also tends to be negative. The mean (median) three-day cumulative size-adjusted return centered on the forecast release date, defined as the firm's compounded raw return minus the compounded return on the size decile portfolio to which the firm belongs at the beginning of the calendar year, is −0.21 percent (−0.17 percent).[11]

Table 1 also provides descriptive statistics on the factors associated with relative forecast accuracy. We base our selection of the factors on prior research; these factors include prior accuracy, forecast age, firm experience, general experience, forecast frequency, number of firms followed, number of industries followed, and brokerage house size (e.g., Clement 1999; Jacob et al. 1999). In addition to the factors considered in prior studies, we also include indicator variables for analyst turnover and award status. We include turnover because Mikhail et al. (1999) find that turnover is negatively related to relative forecast accuracy. We include award status because Stickel (1992) documents that analysts named to the *Institutional Investor* All-American team issue more accurate forecasts.

The forecast revisions are issued an average of 49.79 calendar days prior to the quarterly earnings announcement date; the median forecast age (*FCAGE*) is 44 calendar days. The sample analysts tend to have followed the firm for about three years; the mean (median) number of prior quarters for which the analyst has issued a quarterly earnings forecast for the firm, *FIRMEXP*, is 13.83 (10.0). The mean (median) level of general experience (*GENEXP*), defined as the number of prior firm-quarters the analyst has issued a quarterly forecast for any firm on the Zacks database, is 327.31 (244.0).[12]

TURNOVER is an indicator variable that equals 1 if the analyst changed brokerage houses during the year, 0 otherwise. The mean value of *TURNOVER* indicates that 12.35 percent of our observations correspond to forecast revisions issued by analysts who changed employers during the year. Examining the frequency of turnover of our sample analysts, we find that approximately 33 percent of our sample analysts change brokerage houses at least once during our sample period.

The sample analysts tend to be active forecasters; the mean (median) number of quarterly forecasts issued by the analyst for a firm per quarter, *FCFREQ*, is 4.68 (4.0). The sample analysts issue quarterly earnings forecasts for an average of 18.06 firms during the quarter (*NOFIRM*) and follow an average of 5.29 different two-digit SIC codes (*NOIND*). *BROKSIZE* measures the size of the brokerage house, defined as the number of analysts issuing earnings forecasts at the brokerage house during the year. The average brokerage house represented in the sample has 41.18 analysts issuing earnings forecasts; the median brokerage house size is 35 analysts. The mean value of *IIAWARD* indicates that 21.61 percent of our sample observations correspond to forecast revisions issued by analysts who were named to the *Institutional Investor* All-American team in the prior year. Examining the frequency of award listing of our sample analysts, we find that approximately 17 percent of our sample analysts were named to the *Institutional Investor* All-American team at least once during our sample period.

VI. EMPIRICAL RESULTS

We first provide the results from using the lens model methodology on the entire sample, including estimation results for the factors associated with relative forecast accuracy and investors' reaction to forecast revisions conditional on these factors. This analysis ensures that the findings in prior research on the factors related to forecast accuracy hold in our sample. Then, we present the results from testing our hypotheses.

---

[11] All empirical results hold if we define CAR as the firm's cumulative return less the cumulative return on the value- or equal-weighted market index.

[12] These measures of experience are calculated from the Zacks database, which begins only in 1980. As a result, the experience measures may be understated, especially for analyst-firm observations in the early years of our sample. All conclusions, however, hold if we estimate our empirical models using observations from 1993 to 1999 or from 1997 to 1999.

Results: Overall Sample

In order to calculate the matching index, we first estimate the following expanded version of Equation (1), which represents the statistical model of relative forecast accuracy:

$$\begin{aligned} PRMAPE_{i,j,q} = {} & \alpha_0 + \alpha_1 RFCAGE_{i,j,q} + \alpha_2 PRMAPE_{i,j,q-1} + \alpha_3 RFIRMEXP_{i,j,q} \\ & + \alpha_4 RGENEXP_{i,j,q} + \alpha_5 RTURNOVER_{i,j,q} + \alpha_6 RFCFREQ_{i,j,q} \\ & + \alpha_7 RNOFIRM_{i,j,q} + \alpha_8 RNOIND_{i,j,q} + \alpha_9 RBROKSIZE_{i,j,q} \\ & + \alpha_{10} RIIAWARD_{i,j,q} + \varepsilon_{i,j,q} \end{aligned} \tag{5}$$

where:

$PRMAPE_{i,j,q}$ = the absolute value of analyst *i*'s forecast error for firm *j* and quarter *q* minus the mean absolute forecast error for firm *j* and quarter *q*, deflated by the mean absolute forecast error for firm *j* and quarter *q*;

$RFCAGE_{i,j,q}$ = the number of calendar days between the issuance of analyst *i*'s forecast for firm *j* and quarter *q* minus the mean forecast age for firm *j* and quarter *q*;

$RFIRMEXP_{i,j,q}$ = the firm experience of analyst *i* for firm *j* at quarter *q* minus the mean firm experience for firm *j* and quarter *q*;

$RGENEXP_{i,j,q}$ = the general experience of analyst *i* at quarter *q* minus the mean general experience for firm *j* and quarter *q*;

$RTURNOVER_{i,j,q}$ = an indicator variable that equals 1 if analyst *i* changed brokerage houses during the year in which quarter *q* falls (0 otherwise) minus the mean of the indicator variable for firm *j* and quarter *q*;

$RFCFREQ_{i,j,q}$ = the number of forecasts issued by analyst *i* for firm *j* and quarter *q* minus the mean forecast frequency for firm *j* and quarter *q*;

$RNOFIRM_{i,j,q}$ = the number of firms followed by analyst *i* at quarter *q* minus the mean number of firms followed for firm *j* and quarter *q*;

$RNOIND_{i,j,q}$ = the number of industries followed by analyst *i* at quarter *q* minus the mean number of industries followed for firm *j* and quarter *q*;

$RBROKSIZE_{i,j,q}$ = the size of the brokerage house at which analyst *i* is employed at quarter *q* minus the mean brokerage house size for firm *j* and quarter *q*;

$RIIAWARD_{i,j,q}$ = an indicator variable that equals 1 if analyst *i* was named to the *Institutional Investor* All-American list during the year prior to that in which quarter *q* falls (0 otherwise) minus the mean of the indicator variable for firm *j* and quarter *q*; and

$\varepsilon_{i,j,q}$ = error term with assumed zero mean and constant variance.

The fitted values from the statistical model, $\widehat{PRMAPE}^{STAT}_{i,j,t}$, are those from Equation (5).

To obtain fitted values from investors' (implied) prediction model of relative forecast accuracy, we separately estimate the following expanded version of Equation (3):

$$\begin{aligned} CAR_{i,j,t} = {} & \beta_0 + \alpha'_0 FCREV_{i,j,t} + \alpha'_1 FCREV_{i,j,t} * RFCAGE_{i,j,q} \\ & + \alpha'_2 FCREV_{i,j,t} * PRMAPE_{i,j,q-1} + \alpha'_3 FCREV_{i,j,t} * RFIRMEXP_{i,j,q} \\ & + \alpha'_4 FCREV_{i,j,t} * RGENEXP_{i,j,q} + \alpha'_5 FCREV_{i,j,t} * RTURNOVER_{i,j,q} \\ & + \alpha'_6 FCREV_{i,j,t} * RFCFREQ_{i,j,q} + \alpha'_7 FCREV_{i,j,t} * RNOFIRM_{i,j,q} \\ & + \alpha'_8 FCREV_{i,j,t} * RNOIND_{i,j,q} + \alpha'_9 FCREV_{i,j,t} * RBROKSIZE_{i,j,q} \\ & + \alpha'_{10} FCREV_{i,j,t} * RIIAWARD_{i,j,q} + \zeta_{i,j,q} \end{aligned} \tag{6}$$

where:

$CAR_{i,j,t}$ = three-day ($t-1$, $t+1$) size-adjusted return surrounding the issue of the revised quarterly earnings forecast by analyst i for firm j at time t;
$FCREV_{i,j,t}$ = the price-deflated revision in analyst i's forecast for firm j issued at time t; and
$\zeta_{i,j,t}$ = error term with assumed zero mean and constant variance.

Investors' fitted values, $\widehat{PRMAPE}^{INV}_{i,j,t}$, are obtained from replacing the estimated coefficients in Equation (5) with the estimated coefficients from Equation (6).[13] As described in Equation (4), the matching index is the correlation between the fitted values from the statistical model (Equation (5)) and the fitted values from investors' implied prediction model (Equation (6)).

The results from estimating Equation (5) on the entire sample are provided in the first two columns of Table 2. Consistent with prior research (e.g., O'Brien 1988), the relative age of the forecast, *RFCAGE*, is positively associated with *PRMAPE*. Since higher levels of *PRMAPE* represent less accurate forecasts, the positive coefficient indicates that forecasts issued earlier tend to be less accurate.

Although relative forecast age is the dominant explanatory variable for relative forecast accuracy, consistent with Brown and Mohammed (2001), the other independent variables do add significant explanatory power beyond that contained in relative forecast age. The estimated coefficient on relative prior accuracy is positive and statistically significant, indicating that forecasts issued by analysts who were less accurate in the past are less accurate in the current period.[14] This finding supports Brown (2001) who documents that prior accuracy is an important determinant of current accuracy.[15]

The results for the two experience variables, *RFIRMEXP* and *RGENEXP*, suggest that forecasts issued by more experienced analysts are more accurate. Inconsistent with Mikhail et al.'s (1999) result that analysts who change brokerage houses tend to be less accurate,

---

[13] In Equation (5), higher levels of *PRMAPE* indicate less accurate forecasts. In Equation (6), we expect that investors react more strongly to forecasts that are predicted to be more accurate. Therefore, we multiply the estimated coefficients in Equation (6) by –1 before substituting them into Equation (5).

[14] Replacing $PRMAPE_{q-1}$ by a variable representing the analyst's historical accuracy for the firm (measured over the prior four or eight quarters) yields virtually identical results.

[15] Despite the importance of prior accuracy and forecast age, the other independent variables add significant explanatory power beyond that contained in relative forecast age and relative prior accuracy. The F-statistic testing the null hypothesis that all independent variables except *RFCAGE* and *PRMAPE* equal zero is significantly different from zero ($F = 20.14$, $p < 0.01$, not tabulated).

TABLE 2
Factors Associated with Relative Forecast Accuracy and the Market Reaction to Analyst Forecast Revisions: Entire Sample (n = 101,238)[a]

Equation (5): Forecast Accuracy		Equation (6): Market Reaction	
Intercept	0.0000 (0.00)	Intercept	−0.0008** (−4.82)
		FCREV	0.8491** (31.75)
RFCAGE	0.0038** (46.71)	**RFCAGE*	0.0004 (0.38)
$PRMAPE_{q-1}$	0.0399** (13.53)	$*PRMAPE_{q-1}$	−0.1645** (−4.39)
RFIRMEXP	−0.0009** (−2.83)	**RFIRMEXP*	0.0081* (1.72)
RGENEXP	−0.0000** (−3.28)	**RGENEXP*	−0.0003 (−1.35)
RTURNOVER	0.0036 (0.51)	**RTURNOVER*	−0.0600 (−0.60)
RFCFREQ	−0.0020* (−2.05)	**RFCFREQ*	0.0758** (5.80)
RNOFIRM	0.0013** (3.48)	**RNOFIRM*	−0.0053 (−0.98)
RNOIND	0.0042** (4.05)	**RNOIND*	−0.0269* (−1.85)
RBROKSIZE	0.0005** (5.36)	**RBROKSIZE*	0.0011 (0.80)
RIIAWARD	−0.0209** (−3.18)	**RIIAWARD*	−0.0845 (−1.01)
Adjusted R^2	2.54%	Adjusted R^2	1.00%
F-statistic[b]	39.45**	F-statistic[b]	7.86**

Matching Index = 0.2739

---

*, ** $p < 0.10$ and $p < 0.01$, respectively, two-tailed.

[a] This table provides the estimated coefficients, t-statistics (in parentheses below the estimated coefficient), and adjusted R^2s from estimating Equations (5) and (6) on the sample of 101,238 quarterly forecast revisions. In Equation (5), the dependent variable is *PRMAPE*, defined as the analyst's absolute forecast error for the firm and quarter, minus the mean absolute forecast error of other analysts following the firm for that quarter, deflated by the mean absolute forecast error. In Equation (6), the dependent variable is *CAR*, defined as the three-day size-adjusted return centered on the forecast release date. The prefix "R" for the independent variables indicates that the mean value of the variable for the firm and quarter is subtracted from the observation's value. Variables are defined in Table 1.

[b] For Equation (5), the F-statistic tests the null hypothesis that all independent variables except *RFCAGE* equal zero. For Equation (6), the F-statistic tests the null hypothesis that all interaction terms equal zero.

the estimated coefficient on *RTURNOVER* is insignificant in the full sample estimations. The estimated coefficient on *RFCFREQ* is significantly negative, indicating that analysts who revise their forecasts more frequently are more accurate. The estimated coefficients on *RNOFIRM* and *RNOIND* are positive and significant, indicating that forecasts issued by analysts who follow more firms or more industries tend to be less accurate. In contrast to prior work, we find that the estimated coefficient on *RBROKSIZE* is significantly positive, indicating that forecasts issued by analysts who are employed by larger brokerage houses tend to be less accurate.[16] Finally, the significantly negative coefficient on *RIIAWARD* is consistent with award-winning analysts issuing more accurate forecasts.[17] Overall, these results indicate that the relations between forecast accuracy and these factors documented in prior research (e.g., Clement 1999; Jacob et al. 1999) generally hold in our sample.

Columns three and four in Table 2 provide the results for estimating Equation (6) on the entire sample. The market reaction surrounding the release of the revised forecast is, as expected, positively associated with the signed magnitude of the forecast revision. The interaction terms of *FCREV* with the determinants of forecast accuracy provide incremental explanatory power for the market reaction beyond that contained in *FCREV*; the null hypothesis that all interaction terms are zero is rejected at two-tailed $p < 0.01$.

The results on whether investors, on average, react to factors associated with relative forecast accuracy in a manner consistent with how these factors are related to accuracy are mixed. For example, investors react less strongly to forecasts issued by analysts who were less accurate in the prior quarter, consistent with the relation of this factor to accuracy. Investors also react more strongly to forecasts issued by analysts who have more firm-specific forecasting experience, who are frequent forecasters, and who follow relatively fewer industries. However, they tend not to react differentially to forecasts of different age, inconsistent with this variable's clear importance in predicting accuracy. Further, the results indicate that investors, on average, do not react differentially to forecast revisions issued by analysts who change employers, by analysts who follow more firms, or by analysts with *Institutional Investor* All-American status.

Based on these estimations, the matching index for the entire sample is 0.2739, which is significantly less than 1 at $p < 0.01$ based on the Fisher Z-statistic (Zar 1999). These findings are consistent with those of Clement and Tse (2003), who reject the null hypothesis that the estimated coefficients on the factors they consider in the statistical model equal the coefficients on their factors in the market reaction model using the Mishkin (1983) technique. In the next section, we present the results of our main investigation into differences in the quality of sophisticated and unsophisticated investors' models for predicting forecast accuracy.

Results: Hypothesis Tests for the Effects of Investor Sophistication

In this section, we first test H1 by examining whether the fitted values from investors' implied prediction model of the relative accuracy of forecast revisions more closely match the fitted values from the statistical model when the marginal investor is more likely to be sophisticated. Second, to test H2, we compare the decrease in the matching index from the model that contains the two high-benefit factors to more comprehensive models for sophisticated investors to the same decrease for unsophisticated investors.

Sophistication Characteristics

Table 3 provides descriptive statistics on the investor sophistication characteristics we examine. We use firm characteristics to proxy for the sophistication of the marginal investor,

---

[16] This finding, however, appears to be concentrated in observations from a few years. When Equation (5) is estimated by year, the coefficient on *RBROKSIZE* is significantly positive in only four of the 19 years.

[17] The results are qualitatively similar if we replace *RIIAWARD* with *RWSJAWARD*, an indicator variable that equals 1 if the analyst was named to the *Wall Street Journal* All-Star Analyst list during the year prior to that in which quarter q falls minus the mean of the indicator variable for firm j and quarter q. We tabulate the results for *RIIAWARD* since data on the membership of the *Institutional Investor* All-American team are available for our entire sample period, whereas data to calculate *RWSJAWARD* are only available beginning in 1993.

TABLE 3
Descriptive Statistics on Investor Sophistication Characteristics

Panel A: Descriptive Statistics

Variable[a]	Mean	Median	First Quartile	Third Quartile	Std. Dev.
ANFOLL	13.24	12.00	8.00	18.00	7.49
INST%	52.59	55.06	38.00	67.31	21.61
INST#	191.82	124.00	53.00	272.00	186.06
INST_SH	59.76	24.81	7.48	66.98	92.30
TRADE_DV	19,993.53	17,062.50	11,250.00	26,025.00	11,630.67

Panel B: Pearson and Spearman Rank Correlations[b]

	ANFOLL	*INST%*	*INST#*	*INST_SH*	*TRADE_DV*
ANFOLL		0.2104 (<0.001)	0.5495 (<0.001)	0.4927 (<0.001)	0.3105 (<0.001)
INST%	0.2533 (<0.001)		0.2241 (<0.001)	0.1565 (<0.001)	0.3362 (<0.001)
INST#	0.6196 (<0.001)	0.3880 (<0.001)		0.8667 (<0.001)	0.2675 (<0.001)
INST_SH	0.6516 (<0.001)	0.4277 (<0.001)	0.9517 (<0.001)		0.1161 (<0.001)
TRADE_DV	0.3222 (<0.001)	0.3773 (<0.001)	0.3679 (<0.001)	0.3147 (<0.001)	

[a] Variable definitions: The number of analysts following the firm (*ANFOLL*), the percentage of the firm's common shares held by institutions (*INST%*), the number of institutions holding shares in the firm (*INST#*), the number of shares held by institutions in millions (*INST_SH*), and the median dollar value of shares traded (*TRADE_DV*).

[b] Pearson (Spearman) correlations are above (below) the diagonal. Two-tailed probability values are in parentheses below the correlations. The sample size for all variables except *TRADE_DV* is 101,238; the sample size for *TRADE_DV* is 59,689.

and focus on those firm characteristics identified in prior empirical studies that are associated with the efficiency of a firm's stock price: analyst following, percentage of stock held by institutions, number of institutions owning shares, number of shares held by institutions, and the dollar value of shares traded (see Hand 1990; Walther 1997; Bartov et al. 2000; Brown and Han 2000; Ayers and Freeman 2001; Bhattacharya 2001; Jiambalvo et al. 2002; Piotroski and Roulstone 2002).[18] Analyst following (*ANFOLL*) is the number of analysts on Zacks issuing quarterly earnings forecasts for the firm in the prior year. The percentage of stock held by institutions (*INST%*) is the number of shares held by institutions from Spectrum divided by the common shares outstanding from Compustat as of the beginning of the year. Data on the number of institutions holding shares (*INST#*) and the number of shares held by institutions (*INST_SH*) for the firm as of the beginning of the year are obtained from Spectrum. We measure trade size (*TRADE_DV*) as the median dollar value of shares traded over the prior year. As shown in Panel A, our sample firms are followed by an average of 13.24 analysts. On average, approximately 192 institutions hold

---

[18] Prior research also has used firm size as a proxy for investor sophistication (e.g., Walther 1997). We do not use firm size in our study because this variable also may reflect differences in information environments between firms we classify as having sophisticated and unsophisticated investors. Instead, we control for the effects of firm size by defining each variable on a relative basis, which controls for firm and time-period effects.

59.76 million, or 53 percent, of the sample firms' outstanding common shares. The average of the dollar value of shares traded for our sample firms is $19,994.

Panel B of Table 3 provides the Pearson (Spearman rank) correlations among the sophistication proxies above (below) the diagonal. As expected, all investor sophistication characteristics are positively correlated, suggesting that these firm variables reflect a similar underlying construct. Therefore, we perform principal factor analysis to identify common factor(s) and to classify observations into groups based on these firm characteristics. This procedure reduces the dimensionality of the data by identifying linear combinations of the individual firm characteristics that explain the shared variance among the characteristics (see Bushee 1998). Since data on the dollar value of shares traded from the Trade and Quote Database are only available beginning in 1993, we perform two principal factor analyses, one using the four analyst following and institutional ownership variables over 1981–1999, and the other using all five variables over 1993–1999.

Panel A of Table 4 presents the results of the factor analysis, which produced one common factor.[19] Regardless of whether dollar value of shares traded is included in the analysis, the one common factor identified explains approximately all of the variability

TABLE 4
Factor Analysis of Investor Sophistication Characteristics[a]

Panel A: Factor Analysis

Variable	1981–1999	1993–1999
ANFOLL	0.5803	0.6096
INST%	0.2390	0.2720
INST#	0.9172	0.9230
INST_SH	0.8845	0.8706
TRADE_DV		0.3260
Variability Explained	105.74%	93.34%

Panel B: Descriptive Statistics Based on Factor Scores

	1981–1999		1993–1999	
Variable	Low Soph.	High Soph.	Low Soph.	High Soph.
Mean *ANFOLL*	7.62	19.03**	7.51	19.83**
Mean *INST%*	37.25	59.79**	36.48	62.49**
Mean *INST#*	40.49	405.81**	39.72	400.03**
Mean *INST_SH*	5.66	145.44**	5.77	155.07**
Mean *TRADE_DV*	14,450.40	24,463.45**	13,313.81	26,002.46**

** Statistically different from the corresponding value for the low sophistication group using a t-test for means at two-tailed $p < 0.01$.

[a] Panel A provides the pattern weights and variability explained from principal factor analysis of the investor sophistication characteristics. The first factor analysis excludes the median dollar value of shares traded (*TRADE_DV*), and is based on 101,238 observations from 1981–1999. The second factor analysis includes *TRADE_DV*, and is based on 59,689 observations from 1993–1999. Panel B provides the mean values of the investor sophistication characteristics for subsamples based on the standardized factor scores. Standardized factor scores in the bottom (top) 33 percent are defined to indicate low (high) investor sophistication. Variables are defined in Table 3.

---

[19] We used three criteria to determine the appropriate number of factors to retain: scree plots, minimum eigenvalues, and the proportion of variability explained. All criteria indicated that only one factor should be kept.

among the characteristics. Based on this analysis, we calculate standardized factor scores and rank our sample observations on the basis of these scores. We then divide the sample observations into three equal-sized groups. We classify observations in the top (bottom) group as high (low) investor sophistication. All results discussed below hold if we place the sample observations into two or four groups, or if we separately classify observations into groups based on the five firm characteristics, rather than the standardized scoring coefficients from the factor analysis. As expected, the mean and median analyst following, institutional ownership, and dollar value of shares traded is statistically higher for the high sophistication group than for the low sophistication group (Panel B, median results not tabulated). Therefore, our assumption that the marginal investor of firms with higher values of the standardized factor scores is more likely to be sophisticated is consistent with prior research that analyst following, institutional ownership, and dollar value of shares traded are positively correlated with investor sophistication.

Tests of Hypothesis 1

To test whether sophisticated investors have greater knowledge overall about how to use the information in factors associated with forecast accuracy, we separately estimate Equations (5) and (6) for the sophisticated and unsophisticated investor groups. We then use these estimations to calculate the matching index for each group. We estimate the forecast accuracy regression (Equation (5)) separately for each group to allow for the possibility that the factors related to forecast accuracy vary cross-sectionally with investor sophistication. We estimate the market reaction regression (Equation (6)) separately for each group to allow investors' implied models of relative forecast accuracy to vary cross-sectionally with investor sophistication.[20] We do not tabulate the results for estimating Equations (5) and (6) separately for these subsamples. Rather, we focus on the matching indexes.

Table 5 presents the matching indexes to test H1. We find that the matching index for the sophisticated group is more than two times that for the unsophisticated group. This finding holds regardless of whether the dollar value of shares traded is included in the factor analysis. For example, in the 1993–1999 analysis, the matching index is 0.3943 for sophisticated investors, compared to 0.1166 for unsophisticated investors. The matching index for the unsophisticated group is significantly lower than the corresponding matching index for the sophisticated group at $p < 0.01$ based on the Fisher Z-statistic to test for differences in correlations (Zar 1999, 387), providing strong support for H1. Untabulated results indicate that unsophisticated investors' weights on some factors, most notably *RFCAGE*, are opposite in sign to the weight in the statistical model. Clement and Tse (2003) find that the coefficient on forecast age becomes insignificant once they control for prior uncertainty, and interpret the incorrect sign on *RFCAGE* as suggesting that investors prefer earlier forecasts due to greater uncertainty. In contrast, we find that sophisticated (unsophisticated) investors' weight on *RFCAGE* is of the same (opposite) sign as that in the statistical model, even after controlling for prior uncertainty (see Section VII). Therefore, an alternative explanation for Clement and Tse's (2003) finding is that unsophisticated investors lack knowledge of the information in *RFCAGE* for forecast accuracy. In sum, we

---

[20] We replicated our tests using year-by-year regressions to better reflect the information available to investors. In this analysis, we estimated the accuracy regression using data from year $t-1$, and the market regression using data from year t. We then assessed how well investors' implied predictions of accuracy in year t corresponded to the historical relation between forecast accuracy and the factors. All inferences are unchanged. Further, the results are unchanged if we compare each group of investors' models to the accuracy models estimated on the overall sample.

TABLE 5
Matching Indexes Conditional on Investor Sophistication[a]

	Factor Score Based on *ANFOLL*, *INST%*, *INST#*, and *INST_SH* (1981–1999)	Factor Score Based on *ANFOLL*, *INST%*, *INST#*, *INST_SH*, and *TRADE_DV* (1993–1999)
Low Sophistication	0.1714	0.1166
High Sophistication	0.3614**	0.3943**

** Statistically greater than the corresponding correlation for the bottom group at two-tailed $p < 0.01$ using the Fisher Z test.

[a] The matching index is the Pearson correlation between the fitted values for forecast accuracy from Equation (5) and the investors' implied predictions from Equation (6). In calculating the matching indexes, Equations (5) and (6) are estimated separately for each subsample formed on the basis of the standardized score from the principal factor analysis of the investor sophistication characteristics (see Table 4). The matching indexes are provided for the low and high sophistication subsamples, defined as observations with scores in the bottom 33 percent and top 33 percent, respectively.

conclude that sophisticated investors have greater knowledge overall than unsophisticated investors about the appropriate signs and weights for the set of factors that can be used to predict forecast accuracy.

Tests of Hypothesis 2

Table 6 presents evidence related to H2, which proposes that sophisticated investors have greater knowledge of the net benefit of individual factors that can be used to predict relative forecast accuracy than do unsophisticated investors. To test this hypothesis, we first assess the benefit of each of the factors in Equation (5) for explaining relative forecast accuracy using stepwise model selection techniques.[21] Using the findings from this analysis, we order the factors based on their contribution to the explanatory power of Equation (5) for relative forecast accuracy. We then estimate variations of Equations (5) and (6). The stepwise model selection analysis indicates that relative forecast age has the greatest partial R^2, followed by relative prior accuracy. These findings are consistent with Brown (2001), who documents that this two-factor model performs as well (in terms of predictive ability and explanatory power) as a more complex model based on multiple factors. Therefore, we use this two-factor model to represent the set of high-benefit factors. Subsequent variations of the model progressively add factors that have decreasing benefits for explaining relative forecast accuracy. If sophisticated investors better understand the benefits of factors, then there will be a larger decrease in their matching index (compared to unsophisticated investors) when moving from the model that includes the two high-benefit factors for forecast accuracy (forecast age and prior accuracy) to models that also include lower benefit factors.

Table 6 presents the matching indexes for the overall sample and for subsamples based on investor sophistication for the various models. The conclusion based on the matching indexes for all the models estimated on the overall sample (Panel A) is consistent with findings discussed in the previous section. That is, regardless of the independent variables included in the model, the matching index is less than 1. This result suggests that investors,

---

[21] Ideally, this analysis would be based on the benefits of each factor net of any acquisition and processing costs. An assessment of the costs relative to the benefits would be highly subjective. Consequently, we limit this analysis to an ordering of the factors based on an objective estimate of their key benefit, that of predicting relative forecast accuracy. As a result, our analysis implicitly assumes that the ordering based on this benefit is identical to that based on net benefit.

TABLE 6
Matching Indexes For Simple and Complex Models: Overall Sample and Conditional on Investor Sophistication[a]

	Variables Included in Model									
	RFCAGE (Model 1)	+ *RPMAPE* (Model 2)	+ *RFIRMEXP* (Model 3)	+ *RNOIND* (Model 4)	+ *RBROKSIZE* (Model 5)	+ *RIIAWARD* (Model 6)	+ *RFCFREQ* (Model 7)	+ *RNOFIRM* (Model 8)	+ *RGENEXP* (Model 9)	+ *RTURNOVER* (Model 10)
Panel A: Overall Sample										
	1.0000	0.3463	0.3797	0.3875	0.3842	0.3549	0.2938	0.2887	0.2751	0.2739
Panel B: Measure of Investor Sophistication: Factor Score Based on ANFOLL, INST%, INST#, and INST_SH (1981–1999)										
Low Soph.	−1.0000	0.1345	0.1951	0.2152	0.2038	0.1736	0.2041	0.2132	0.1730	0.1714
			0.0606	*0.0807*	*0.0693*	*0.0391*	*0.0696*	*0.0787*	*0.0385*	*0.0369*
High Soph.	1.0000**	0.9916**	0.9736**	0.6235**	0.6261**	0.5406**	0.4271**	0.3798**	0.3613**	0.3614**
			−0.0180	***−0.3681***	***−0.3655***	***−0.4510***	***−0.5645***	***−0.6118***	***−0.6303***	***−0.6302***
Panel C: Measure of Investor Sophistication: Factor Score Based on ANFOLL, INST%, INST#, INST_SH, and TRADE_DV (1993–1999)										
Low Soph.	−1.0000	−0.0311	0.0545	0.0868	0.0616	0.0632	0.0986	0.1130	0.1006	0.1166
			0.0856	*0.1179*	*0.0927*	*0.0943*	*0.1297*	*0.1441*	*0.1317*	*0.1477*
High Soph.	1.0000**	0.9928**	0.9853**	0.8468**	0.8457**	0.6930**	0.4306**	0.4090**	0.4030**	0.3943**
			−0.0075	***−0.1460***	***−0.1471***	***−0.2998***	***−0.5622***	***−0.5838***	***−0.5898***	***−0.5985***

** Statistically greater than the corresponding correlation for the low sophistication subsample at two-tailed $p < 0.01$ using the Fisher Z test.

[a] This table provides the matching indexes for various models for the overall sample (Panel A) and subsamples formed on the basis of the standardized score from the principal factor analysis of the investor sophistication characteristics (see Table 4). Within Panels B and C, the indexes are provided for the low and high sophistication subsamples, defined as observations with scores in the bottom 33 percent and top 33 percent, respectively. The difference relative to Model 2 is provided in italics under the matching index. The difference for the high sophistication subsample is in bold when it is statistically smaller than the corresponding difference for the low sophistication subsample at two-tailed $p < 0.01$ using a Tukey-type test for differences in correlations.

on average, use some (but not all) of the information in these factors when predicting relative forecast accuracy.

Panels B and C of Table 6 present the results conditional on investor sophistication. Consistent with the findings provided in Table 5 to test H1, the matching index for the unsophisticated subsample is significantly lower than for the sophisticated subsample using a Fisher Z-statistic, regardless of the independent variables included. This finding holds regardless of whether the dollar value of shares traded is included in the factor analysis. Consistent with the untabulated results discussed above, the matching index for Model 1 in Panels B and C is -1 for unsophisticated investors, indicating that unsophisticated investors' weight on *RFCAGE* is of opposite sign to that in the statistical model.[22] Overall, our findings supporting H1 hold regardless of the factors included.

Table 6 also provides the change in the matching index (in italics) from Model 2 to Models 3 through 10, respectively, for each subsample. To test H2, we compare the change in the matching index for the sophisticated investor subsample to the change for the unsophisticated investor subsample as factors with decreasing benefits are progressively added as independent variables. In both Panels B and C, we find a consistently larger decrease in the matching index for the subsample in which the marginal investor is likely to be sophisticated. The decrease in the matching index is statistically greater for sophisticated investors than for unsophisticated investors using a Tukey-type test for differences in correlations (Zar 1999, 393). These results are consistent with H2, and suggest that sophisticated investors better understand the benefits of the individual factors that can be used to predict forecast accuracy.

The results in Table 6 indicate that the fitted values derived from sophisticated investors' model correspond less with the fitted values derived from the statistical model when low-benefit factors are included. While consistent with H2, we cannot determine based on these results whether sophisticated investors underweight low-benefit factors more than unsophisticated investors. To investigate this specific form of adaptive decision making, which underlies our prediction in H2, we must compare the coefficients on individual factors in the statistical model to the implied coefficients on those factors from sophisticated and unsophisticated investors' models. We define underweighting to occur when the implied weight sophisticated investors place on a factor is either insignificantly different from zero or less than the weight on that factor in the statistical model.[23] This analysis indicates that sophisticated investors underweight the following low-benefit factors: *RFIRMEXP*, *RGENEXP*, *RTURNOVER*, *RFCFREQ*, *RNOFIRM*, *RBROKSIZE*, and *RIIAWARD*. By contrast, unsophisticated investors overweight *RFIRMEXP*, *RGENEXP*, and *RFCFREQ*. These findings provide additional support for H2, and suggest that sophisticated investors' model better reflects the cost-benefit trade-offs for the factors associated with forecast accuracy than does that of unsophisticated investors.

VII. ROBUSTNESS CHECKS

In this section, we discuss the assumptions underlying our empirical tests and conclusions based on those tests. Where empirically feasible, we provide evidence on the validity of each of these assumptions.

---

[22] This finding holds if we include a control for prior uncertainty in the model (see Section VII).

[23] Regression diagnostics indicate no evidence of the statistical problems of multicollinearity in either the statistical or the investor model, suggesting that the individual coefficient estimates upon which we make these inferences are not biased.

Assumption 1: Sophisticated and Unsophisticated Investors Attempt to Predict Forecast Revision Accuracy

We assume that both sophisticated and unsophisticated investors attempt to predict the accuracy of analysts' forecast revisions, and, in particular, that sophisticated and unsophisticated investors have equivalent incentives for predicting the accuracy of analysts' forecasts. We know of no way to measure directly investors' incentives for predicting accuracy. However, we would expect that, if sophisticated investors had greater incentives for predicting forecast accuracy (e.g., greater expected trading profits), then the correlation of *CAR* and *ex post* accuracy would be greater for sophisticated investors. Therefore, to investigate the validity of this assumption, we calculate the correlation of *CAR* and *ex post* accuracy of forecast revisions separately for our sophisticated and unsophisticated investor subsamples. We are unable to reject the null hypothesis that the correlation is equal across these two subsamples (two-tailed, $p > 0.10$). Additionally, if sophisticated investors face greater incentives to predict forecast accuracy, then we would not find in our tests of H2 that they ignore or underweight less beneficial cues more than do unsophisticated investors. These findings are consistent with both sophisticated and unsophisticated investors attempting to predict the accuracy of analysts' forecast revisions. However, we are unable to completely rule out differences in the incentives for predicting forecast revision accuracy across sophisticated and unsophisticated investors.

Assumption 2: Sophisticated and Unsophisticated Investors Predict Forecast Revision Accuracy Using the Same Set of Observable Factors

Our tests also assume that both sophisticated and unsophisticated investors predict forecast revision accuracy using the set of observable factors that we include in our models. In other words, we must assume that both groups of investors have access to information about all these factors. Further, we must assume that these factors predict forecast accuracy equally well in the sophisticated and unsophisticated subsamples. The first part of this assumption is supported by the similarity of the explanatory power of the market reaction regression (Equation (6)) across the two subsamples.[24] To investigate the second part of this assumption, we examine whether the factors we consider explain relative forecast accuracy equally well in the sophisticated and unsophisticated subsamples. We find that the explanatory power (both in-sample and out-of-sample) of the accuracy regression (Equation (5)) is similar across the two subsamples.[25] Overall, these results suggest that our results are not due to differences between sophisticated and unsophisticated investors as to the factors they use to predict forecast accuracy or to the predictive ability of observable factors.

Assumption 3: Sophisticated and Unsophisticated Investors Obtain Forecast Revisions on a Timely Basis

Our tests assume that unsophisticated and sophisticated investors have equally timely access to forecast revisions. It is possible that the higher matching index for sophisticated investors reflects greater or more timely access to analyst forecast revisions. To investigate this possibility, we perform three tests. First, we investigate whether the explanatory power of *FCREV* for *CAR* is higher in the sophisticated investor subsample. We regress *CAR* on *FCREV* separately for the sophisticated and unsophisticated subsamples and find no evidence that this is the case. Further, in a regression of *CAR* on an intercept, an indicator variable for sophistication ($SOPH = 0$ if unsophisticated, 1 if sophisticated), *FCREV*, and *FCREV*SOPH*, we find that the estimated coefficient on *FCREV*SOPH* is insignificant, suggesting that sophisticated investors do not react more strongly to analyst forecast revisions.[26]

---

[24] As an alternative test of the assumption of equal access to information on the factors that can be used to predict forecast revision accuracy, we classify *FIRMEXP*, *GENEXP*, *FCFREQ*, *NOIND*, and *BROKSIZE* as "not equally accessible," and eliminate them from our models. All of our results for H1 and H2 hold when these "inaccessible" factors are excluded from the analyses. We thank the referee for this suggestion.

[25] We assess the out-of-sample predictive ability of these factors by estimating the accuracy regression using data from year $t-1$, and predicting relative forecast accuracy in year t based on the estimated coefficients from year $t-1$. The absolute (squared) difference between the actual relative accuracy in year t and the predicted relative accuracy based on the year $t-1$ regression for the unsophisticated subsample was either lower than or equal to the absolute (squared) difference for the sophisticated subsample.

Second, we replicate our analyses in Tables 5 and 6 using five- and seven-day returns accumulation windows ($(-3, +3)$ or $(-1, +3)$, where $t = 0$ is the date the revised forecast is issued) to ensure that we fully capture the reaction of all investors. We extend the window to three days after the forecast release date because prior work suggests that virtually all of abnormal trading responses for all investor types occur within three days after announcements (Cready 1988; Lee 1992; Bhattacharya 2001). We extend the window to three days before the forecast release date in case sophisticated investors have access to this information in advance of the formal release date.[27] Third, we replicate our primary analyses for H1 and H2 using observations only from 1997 forward. We select this time period because many investor advice sites and analysts' forecasts became available on the Internet starting in 1997 (Kurtz 2000). All results continue to hold using either the expanded returns accumulation window and/or the more recent time period. Overall, these analyses do not suggest that sophisticated investors have greater or timelier access to analyst forecast revisions.

Assumption 4: Sophisticated and Unsophisticated Investors Update Their Beliefs Based on Forecast Revisions and Predicted Accuracy

Our tests assume that investors use the revised analyst forecast as well as their prediction regarding its accuracy in revising their beliefs about future earnings. Previous research has shown that sophisticated and unsophisticated investors may differ in the extent to which they form earnings expectations on the basis of analysts' forecasts versus seasonal random walk forecasts (Walther 1997; Bhattacharya 2001). Consequently, to ensure that our use of the analyst's prior forecast as the expectation does not drive our results, we replicate our analyses for H1 and H2 using either the seasonal random walk forecast or the consensus analyst forecast as the expectation.[28] These analyses suggest that our assumption that both sophisticated and unsophisticated investors use the analyst forecast revision to update their beliefs does not confound our results.

---

[26] We do not find any evidence that differences in information environments related to firm size are confounded with differences in investor sophistication and, thus, create these findings. The explanatory power of *FCREV* for *CAR* remains similar even after we partition the sample based on firm size into three groups and re-estimate that regression. The coefficient on *FCREV*SOPH* remains insignificant when we include firm size (*SIZE*) and *FCREV*SIZE* in the regression.

[27] All results for H1 and H2 also hold using a $(-7, +7)$ window.

[28] We measure the outstanding consensus forecast as the mean of all forecasts issued after the beginning of the quarter but before the revised forecast is released. We require a minimum of two forecasts to construct the consensus. Our results hold if we define the outstanding consensus forecast as the median of all forecasts issued during this window.

Assumption 5: Sophisticated and Unsophisticated Investors Face a Similar Level of Information Uncertainty

Our tests implicitly assume that sophisticated and unsophisticated investors face a similar level of information uncertainty about upcoming earnings. To investigate this assumption, we compare the mean squared forecast error (MSFE) across the sophisticated and unsophisticated subsamples. Following Clement and Tse (2003), we calculate the average of the squared price-deflated forecast error for all prior quarterly earnings forecasts outstanding at the date of each forecast revision (see also Barron et al. 1998). While the median MSFE is 0.0000 in both subsamples, the mean is significantly higher in the unsophisticated sample (0.0004) than in the sophisticated subsample (0.0002, $p < 0.01$). This finding provides some evidence that unsophisticated investors face a higher level of information uncertainty prior to the forecast revision date. To ensure that our results are not sensitive to our assumption, we replicate our analyses for H1 and H2 after including the mean-adjusted MSFE in our models. All results hold, suggesting that differences in information uncertainty between sophisticated and unsophisticated investors do not cause our findings.

Assumption 6: Sophisticated and Unsophisticated Investors' Stock-Holding Decisions are Based on Revised Earnings Expectations

We assume that individual investors' decisions about whether to buy or sell a stock are based on their revised expectations of future earnings. However, these decisions also could be affected by their risk preferences, the expected benefits (net of costs) of changing holdings, and the incentives they face. For example, sophisticated investors may face greater incentives to justify their investment decisions than do unsophisticated investors. Therefore, sophisticated investors may react more strongly, for example, to forecast revisions issued by analysts named to the *Institutional Investor* All-American team or by analysts employed at larger brokerage houses, not because these factors are associated with the accuracy of the revised forecast, but because these factors provide an easy justification to outside parties or superiors. To examine whether sophisticated investors react more strongly than unsophisticated investors to forecast revisions that would be easier to justify, we regress *CAR* on an intercept, an indicator variable for sophistication (*SOPH* = 0 if unsophisticated, 1 if sophisticated), *FCREV*, *FCREV*SOPH*, and *FCREV*SOPH*RIIAWARD* (or *FCREV*SOPH*RBROKSIZE*). The estimated coefficient on *FCREV*SOPH*RIIAWARD* (*FCREV*SOPH*RBROKSIZE*) is not significant, indicating that sophisticated investors do not react more strongly to forecast revisions issued by analysts who were members of the *Institutional Investor* All-American team (who work for larger brokerage houses). While these results provide some evidence that differences in accountability or incentives are not driving the higher matching index for sophisticated investors, we are unable to rule out that differences in risk preferences, the expected benefits of changing holdings, or other incentives could be an alternative explanation for our results.

Overall, our additional analyses suggest that the conclusions we draw regarding differences between sophisticated and unsophisticated investors' models for predicting the relative accuracy of analysts' forecast revisions are not confounded by the assumptions underlying our empirical tests.

VIII. CONCLUSIONS

In this study we examine the extent to which sophisticated and unsophisticated investors' reactions to sell-side analysts' forecast revisions appear consistent with factors shown to be associated with forecast accuracy. We infer investors' model for predicting the relative accuracy of forecast revisions from their reactions to those revisions, and investigate two differences in the quality of sophisticated and unsophisticated investors' models. We base our inferences regarding differences in the quality of prediction models on the "matching index" from Brunswik's lens model of human judgment.

As predicted, we find that the matching index for the subsample in which the marginal investor is likely to be sophisticated is statistically higher than the matching index for the subsample in which the marginal investor is likely to be unsophisticated. We also find that the decrease in the matching index for a model that includes the two most beneficial factors to those for more comprehensive models is greater for the sophisticated investor subsample than for the unsophisticated investor subsample. Taken together, our results suggest that sophisticated investors not only have more knowledge overall about the factors related to analysts' forecast accuracy, but also that they are aware of the individual factors that are the most beneficial to use.

The findings of this study contribute to the literature in several ways. First, we investigate how different types of investors use an important source of accounting information provided by a key information intermediary—the sell-side analyst's quarterly earnings forecast revision. Consequently, we add to both the growing literature on behavioral differences between sophisticated and unsophisticated investors (e.g., Lee et al. 1999; Shapira and Venezia 2001) and the literature on investors' reactions to analysts' forecasts. Further, our factor analysis of various sophistication proxies may be useful to future research examining investor sophistication issues.

Additionally, we introduce the lens model to the archival accounting literature. The lens model methodology might be a fruitful approach for researchers to consider when investigating the quality of investors' predictions about future outcomes in situations in which the information signals are correlated, e.g., factors related to corporate bankruptcy. Further, we provide evidence on the validity of the assumptions underlying our application of the lens model. These analyses should be useful to future researchers who seek to make inferences about individual investor behavior using market-level data.

Our study has both benefits and limitations related to the archival market-level approach we employ. The most salient benefit is that we demonstrate differences between investors we classify as sophisticated and unsophisticated in a market setting, with its attendant incentives and disciplinary mechanisms. Further, we demonstrate that both groups of investors appear to be predicting the accuracy of analysts' forecasts in this setting. Although we are able to provide some empirical evidence on the validity of most of the assumptions necessary for inferring investors' prediction models from market reactions, we cannot examine all of them. In particular, we cannot investigate sophistication-related differences in risk preferences and other such factors. As a result, we are unable to rule out all alternative explanations for our findings. Future experimental research could attempt to replicate our findings in a setting that would control for differences between groups of investors. Finally, experimental work that probes the processes that underlie investors' development of models to predict forecast accuracy ultimately would be necessary to provide the basis for specific suggestions for investor education programs (Bonner 1999).

REFERENCES

Ackerman, P., R. Sternberg, and R. Glaser. 1989. *Learning and Individual Differences: Advances in Theory and Research*. New York, NY: W.H. Freeman and Co.

Ashton, R. 1982. *Human Information Processing in Accounting*. Sarasota, FL: American Accounting Association.

Ayers, B., and R. Freeman. 2001. Evidence that price leads of earnings increase with analyst following and institutional ownership. Working paper, University of Georgia.

Barber, B., and T. Odean. 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* (April): 773–806.

Barron, O., O. Kim, S. Lim, and D. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* (October): 421–433.

Bartov, E., S. Radhakrishnan, and I. Krinsky. 2000. Investor sophistication and patterns in stock returns after earnings announcements. *The Accounting Review* (January): 43–63.

Beaver, W. 1998. *Financial Reporting: An Accounting Revolution.* 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Bettman, J., E. Johnson, and J. Payne. 1990. A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes* (February): 111–139.

Bhattacharya, N. 2001. Investors' trade size and trading responses around earnings announcements: An empirical investigation. *The Accounting Review* (April): 221–244.

Bonner, S. 1990. Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review* (January): 72–92.

———, and P. Walker. 1994. The effects of instruction and experience on the acquisition of auditing knowledge. *The Accounting Review* (January): 157–178.

———. 1999. Judgment and decision-making research in accounting. *Accounting Horizons* (December): 385–398.

Brous, P., and O. Kini. 1994. The valuation effects of equity issues and the level of institutional ownership: Evidence from analysts' earnings forecasts. *Financial Management* (Spring): 33–46.

Brown, L., and J. Han. 2000. Do stock prices fully reflect the implications of current earnings for future earnings for AR1 firms? *Journal of Accounting Research* (Spring): 149–164.

———. 2001. How important is past analyst earnings forecast accuracy? *Financial Analysts Journal* (November/December): 44–49.

———, and E. Mohammad. 2001. Profiting from predicting individual analyst earnings forecast accuracy. Working paper, Georgia State University.

Brown, P., G. Foster, and E. Noreen. 1985. *Security Analyst Multi-Year Earnings Forecasts and the Capital Markets.* Sarasota, FL: American Accounting Association.

Bushee, B. 1998. The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review* (July): 305–333.

Chen, Q., J. Francis, and W. Jiang. 2001. Investor learning about analyst predictive ability. Working paper, Duke University.

Clement, M. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* (June): 285–303.

———, and S. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* (January): 227–249.

Cready, W. 1988. Information value and investor wealth: The case of earnings announcements. *Journal of Accounting Research* (Spring): 1–27.

Einhorn, H. 1974. Expert judgment: Some necessary conditions and an example. *Journal of Applied Psychology* (October): 562–571.

El-Gazzar, S. 1998. Predisclosure information and institutional ownership: A cross-sectional examination of market revaluations during earnings announcement periods. *The Accounting Review* (January): 119–129.

Fennema, M., and D. Kleinmuntz. 1995. Anticipations of effort and accuracy in multiattribute choice. *Organizational Behavior and Human Decision Processes* (July): 21–32.

Givoly, D., and J. Lakonishok. 1979. The information content of financial analysts' earnings forecasts. *Journal of Accounting and Economics* (March): 165–185.

Gleason, C., and C. Lee. 2003. Analyst forecast revisions and market price discovery. *The Accounting Review* (January): 193–225.

Gonedes, N., N. Dopuch, and S. Penman. 1976. Disclosure rules, information production, and capital market equilibrium: The case of forecast disclosure rules. *Journal of Accounting Research* (Spring): 89–137.

Hammond, K., and T. Stewart. 2001. *The Essential Brunswik: Beginnings, Extensions, Applications.* New York, NY: Oxford University Press.

Hand, J. 1990. A test of the extended functional fixation hypothesis. *The Accounting Review* (October): 764–780.

Hassel, C., and M. Norman. 1992. Financial characteristics of neglected and institutionally held stocks. *Journal of Accounting, Auditing, and Finance* (April): 313–330.

Hodge, F. 2001. Investor habits. Speech given at the Annual Meeting of the Puget Sound Chapter of the National Association of Investment Clubs, Women's University Club, Seattle, WA, June 14.

Jacob, J., T. Lys, and M. Neale. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* (November): 51–82.

Jiambalvo, J., S. Rajgopal, and M. Venkatachalam. 2002. Institutional ownership and the extent to which stock prices reflect future earnings. *Contemporary Accounting Research* (Spring): 117–145.

Johnson, E. 1988. Expertise and decision under uncertainty: Performance and process. In *The Nature of Expertise*, edited by M. Chi, R. Glaser, and M. Farr. Hillsdale, NJ: Lawrence Erlbaum.

Johnson, P., A. Duran, F. Hassebrock, J. Moller, M. Prietula, P. Feltovich, and D. Swanson. 1981. Expertise and error in diagnostic reasoning. *Cognitive Science* (September): 235–283.

Kurtz, H. 2000. *The Fortune Tellers: Inside Wall Street's Game of Money, Media, and Manipulation.* New York, NY: The Free Press.

Lee, C. 1992. Earnings news and small traders: An intraday analysis. *Journal of Accounting and Economics* (June/September): 265–302.

Lee, Y., J. Lin, and Y. Liu. 1999. Trading patterns of big versus small players in an emerging market: An empirical analysis. *Journal of Banking and Finance* (May): 701–725.

Lev, B. 1988. Toward a theory of equitable and efficient accounting policy. *The Accounting Review* (January): 1–22.

Maines, L. 1996. An experimental examination of subjective forecast combination. *International Journal of Forecasting* (June): 223–233.

Mikhail, M., B. Walther, and R. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* (Supplement): 131–157.

———, ———, and ———. 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* (April): 185–200.

Mishkin, F. 1983. *A Rational Expectations Approach to Macroeconomics: Testing Policy Ineffectiveness and Efficient-Markets Models.* Chicago, IL: University of Chicago Press.

O'Brien, P. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* (January): 53–83.

———. 1990. Forecast accuracy of individual analysts in nine industries. *Journal of Accounting Research* (Autumn): 286–304.

Park, C., and E. Stice. 2000. Analyst forecasting ability and the stock price reaction to forecast revisions. *Review of Accounting Studies* (September): 259–272.

Payne, J., J. Bettman, and E. Johnson. 1993. *The Adaptive Decision Maker.* New York, NY: Cambridge University Press.

Piotroski, J., and D. Roulstone. 2002. Informed traders and the incorporation of industry and firm-specific information into stock prices. Working paper, University of Chicago.

Potter, G. 1992. Accounting earnings announcements, institutional investor concentration, and common stock returns. *Journal of Accounting Research* (Spring): 146–155.

Richardson, S., S. Teoh, and P. Wysocki. 2001. The walkdown to beatable analyst forecasts: The roles of equity issuance and insider trading incentives. Working paper, University of Michigan.

Shapira, Z., and I. Venezia. 2001. Patterns of behavior of professionally managed and independent investors. *Journal of Banking and Finance* (August): 1573–1587.

Slovic, P. 1969. Analyzing the expert judge: A descriptive study of a stockbroker's decision processes. *Journal of Applied Psychology* (August): 255–263.

SRI International. 1987. *Investor Information Needs and the Annual Report*. Morristown, NJ: Financial Executives Research Foundation.

Stickel, S. 1992. Reputation and performance among security analysts. *Journal of Finance* (December): 1811–1836.

Walther, B. 1997. Investor sophistication and market earnings expectations. *Journal of Accounting Research* (Autumn): 157–179.

Williams, P., G. Moyes, and K. Park. 1996. Factors affecting earnings forecast revisions for the buy-side and sell-side analyst. *Accounting Horizons* (September): 112–121.

Wilson, R. 1975. Informational economies of scale. *Bell Journal of Economics* (Spring): 184–195.

Yunker, J., and T. Krehbiel. 1988. Investment analysis by the individual investor. *Quarterly Review of Economics and Business* (Winter): 90–101.

Zar, J. 1999. *Biostatistical Analysis*. 4th edition. Upper Saddle River, NJ: Prentice Hall.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 707–724

Earnings Predictability and the Direction of Analysts' Earnings Forecast Errors

Michael J. Eames
Santa Clara University

Steven M. Glover
Brigham Young University

ABSTRACT: Das et al. (1998) suggest that as earnings become less predictable, analysts issue increasingly optimistic forecasts to please managers and consequently gain, or at least limit the loss of, access to managers' private information. We reexamine the association between earnings forecast error and earnings predictability because there is evidence suggesting that deliberate earnings forecast optimism is not an effective mechanism for gaining access to managers' information (e.g., Eames et al. 2002; Matsumoto 2002). We document associations between earnings level and both forecast error and earnings predictability. These associations suggest that earnings level may be an important control variable when examining the association between forecast error and earnings predictability. When we control for the level of earnings we find no significant association between forecast error and earnings predictability. Thus, we find no evidence that analysts intentionally issue optimistically biased earnings forecasts.

Keywords: *analyst earnings forecasts; forecast error; forecast bias; earnings predictability.*

Data Availability: *Data are commercially available from the sources identified in the text.*

I. INTRODUCTION

Das et al. (1998) (hereafter DLS) examine the association between Value Line analysts' earnings forecasts and earnings predictability and find that as earnings become less predictable analysts' earnings forecasts become increasingly optimistic. To explain this pattern, DLS extend the management relations hypothesis (Francis and Philbrick 1993) (hereafter FP) to the earnings predictability context and argue that analysts intentionally bias their forecasts to curry favor with management in order to gain access to

---

We are grateful for the comments of Sandra Chamberlain, Suzanne Luttman, the editor, two anonymous reviewers, participants at the 2000 Western Regional Meeting of the AAA and the 2002 European Accounting Conference. We particularly thank Linda Bamber for her insightful recommendations. We thank the PricewaterhouseCoopers Foundation, the Ernst & Young Faculty Research Fund, and the Accounting Development Funds at Santa Clara University and Brigham Young University for financial support. We thank I/B/E/S and Value Line for making their data available.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted May 2002
Accepted January 2003

management's information.[1] DLS posit that as earnings become less predictable, analysts issue increasingly optimistic earnings forecasts because the potential gain in forecast accuracy from access to managers' nonpublic information is greater when earnings are less predictable. This intentional bias improves the overall accuracy of forecasts when the gain in accuracy through access to private information outweighs the loss in accuracy from the optimistic bias. DLS's evidence that analysts intentionally issue optimistically biased earnings forecasts is of interest to regulators (interested in establishing an environment where information is fairly and efficiently transferred to all investors), investors (as they interpret earnings forecast for their investment decisions), firms (as managers try to meet or beat the earnings forecast), and forecasting agencies (as they face the scrutiny of regulators and investors).

Contrary to DLS's theory, recent research suggests that issuing intentionally optimistic earnings forecasts is not an effective means for pleasing managers and improving access to their private information. Jorge and Rees (2000) present interview and survey evidence from Spain that managers more commonly respond to unfavorable analyst reports by providing more, rather than less, information. In addition, optimistic earnings forecasts result in negative earnings surprises, and related negative market reactions, while accurate and pessimistic forecasts have been linked with positive market responses (Bartov et al. 2002; Kinney et al. 2002; Kasznik and McNichols 2002; Lopez and Rees 2002; Skinner and Sloan 2001). Burgstahler and Eames (2002) present evidence that managers avoid negative earnings surprises by both managing earnings upward and analysts' forecasts downward—results inconsistent with a managerial preference for forecast optimism. Matsumoto (2002) argues that managers prefer pessimistic forecasts to avoid adverse market reactions to negative earnings surprises, and presents results consistent with this position.

FP and Kim and Lustgarten (1998) report evidence of greater earnings forecast optimism in association with less favorable *recommendations*, and reason that forecast optimism repairs the damage to the manager-analyst relationship invoked by unfavorable recommendations. However, Eames and Glover (2002) and Eames et al. (2002) further examine the association and conclude that FP's and Kim and Lustgarten's (1998) results likely stem from omitting the correlated variable—earnings level—from their analyses. Without including earnings level in their analyses, Eames and Glover (2002) and Eames et al. (2002) replicate the FP and Kim and Lustgarten (1998) results. However, when they control for earnings level they find no association between forecast optimism and recommendations for Value Line analysts (Eames and Glover 2002) and greater forecast optimism in association with more favorable recommendations for broker analysts (Eames et al. 2002).[2] These findings, combined with recent research indicating that issuing intentionally optimistic earnings forecasts is not an effective method to curry favor with management, lead us to explore whether the level of earnings is also an important control variable when considering the relation between analyst forecast errors and *earnings predictability*. Specifically, we investigate whether DLS's results are driven by their failure to consider the level of earnings in their analyses.

---

[1] The management relations hypothesis (Francis and Philbrick 1993) proposes that analysts intentionally issue earnings forecasts that exceed their true expectations in order to counter the negative impact of unfavorable recommendations and to curry favor with managers and consequently gain, or at least limit the loss of, access to their private information. DLS note anecdotal evidence that managers may limit contact with analysts who are considered negative toward the company (see Drucker and Sapsford [2002] for a recent example of such anecdotal evidence).

[2] Eames et al. (2002) interpret their observation of greater forecast optimism for buy recommendations and greater forecast pessimism for sell recommendations as consistent with their motivated reasoning hypothesis, as well as broker-analysts engaging in trade-boosting behavior.

For earnings level to be an important control variable in examinations of the association between forecast error and earnings predictability, there must be associations between earnings level and both forecast error and earnings predictability. Numerous studies report an inverse relation between forecast error and the level of reported earnings (e.g., Brown 2001; Eames et al. 2002; Eames and Glover 2002; Hwang et al. 1996). The association reflects both earnings shocks due to unanticipated events and earnings management. Unexpected positive (negative) earnings shocks result in higher (lower) earnings and will generally be associated with more forecast pessimism (optimism) (Eames et al. 2001). Earnings baths are associated with low earnings and forecast optimism, while management to beat the forecast will lead to higher earnings and greater forecast pessimism (Abarbanell and Lehavy 2002).

In addition to the association between forecast error and earnings level, we document an association between earnings level and earnings predictability. We find earnings predictability is highest for earnings observations near the cross-sectional median level of earnings and that earnings predictability declines dramatically as earnings increasingly deviates from the median. With an association between forecast error and earnings level, and between earnings level and earnings predictability, we anticipate that earnings level will be an important control variable in examining the association between forecast error and earnings predictability.

We replicate DLS's results using analyses and data similar to theirs. However, plots of earnings predictability and the level of earnings reveal that the association between these two variables differs above and below median earnings. When we partition our sample at median earnings and further control for the level of earnings by including earnings as an independent variable, we find that neither subsample yields significant evidence of an association between forecast optimism and earnings predictability.[3] These findings suggest that DLS's results indicating intentional analyst bias are spurious and attributable to a failure to adequately consider earnings level controls in their analyses.

This study makes several important contributions to the existing literature examining earnings forecast errors. First, we provide evidence that earnings level is an important control variable to consider when examining the association between forecast error and earnings predictability. This finding, in combination with studies demonstrating that earnings is an important control variable when examining the association between forecast error and recommendations (e.g., Eames et al. 2002), suggests that research on earnings forecast error in other contexts should consider controlling for the level of earnings. Second, we demonstrate that forecast-error research that assumes linear relationships can be seriously flawed. A simple plot of the underlying data reveals that DLS's theory and findings do not generalize to at least half of the firms (i.e., those with earnings above the median). Third, we demonstrate the appropriate "switch point" (from optimism to pessimism) in earnings-forecast-error research is the economy-wide earnings mean/median and not zero (as suggested by Brown [2001]). Fourth, we find no significant association between forecast errors and earnings predictability when we include a control for earnings in our model and we conclude that DLS's results are spurious.

---

[3] Partitioning the sample at mean earnings (as well as other points in that vicinity, see footnote 11) yields essentially identical results to partitioning at the median.

II. DATA AND SAMPLE SELECTION

For consistency with DLS, we use Value Line (VL) annual earnings forecasts and realizations. The primary advantage of using VL data is that VL offers neither brokerage nor underwriting services, so VL forecast errors are not confounded by incentives for promoting security transactions. Each week VL issues annual EPS forecasts sequentially covering approximately one-thirteenth of the roughly 1,700 firms followed in the VL *Standard Edition*.[4] Thus in each quarter, VL issues a single annual EPS forecast for each firm followed. From the *VL Estimates and Projections File*, we obtain realized annual EPS and realized shares outstanding. We also obtain forecasts of annual EPS and shares outstanding, as well as share price, the Value Line Timeliness Rank, and the Value Line ValuGauge Relative Position measure, all as of the forecast date. The ValuGauge Relative Position measure provides VL's index of forecasting difficulty. Values range from 1 to nearly 1,600, with low numbers for earnings that are easy to forecast and high numbers for earnings that are difficult to forecast. We limit our sample to forecasts of annual earnings for December fiscal-year-end firms, issued over the four quarters preceding year-end, and obtain a sample of 29,888 forecasts of annual earnings.

To obtain annual forecasted and actual earnings we multiply forecasted EPS and actual EPS by forecasted and actual shares, respectively.[5] Because we obtain actual and forecasted earnings for a broad range of firm sizes, and for consistency with prior forecast error research (DLS; Eames et al. 2002; FP), we scale both earnings and forecast error by the market value of equity at the first fiscal quarter's forecast date, computed as the product of VL reported share price at the first quarter forecast date and number of shares outstanding at the prior year end.[6] We compute annual forecast error as actual earnings less forecasted earnings. To limit the affect of outliers, in those instances where we are not using the DLS method of analyzing average values over the period 1989 to 1993, we delete 456 firm-quarter forecast observations for which the absolute value of earnings exceeds 50 percent of market value or the absolute value of forecast error exceeds 20 percent of market value, resulting in a sample of 29,432 observations.

Table 1, Panel A, presents sample observations by year. Sample sizes for individual years range from a low of 1,639 observations in 1987 to a high of 2,832 observations in 1996. A total of 1,335 firms are represented in the sample. The average number of observations per firm is 20.05. The maximum and minimum number of observations for a firm are 52 and 1, respectively. Panel B of Table 1 provides an industry breakdown using two-digit SIC codes for our sample of 29,432 firm-quarter forecasts of annual earnings.

Given systematic differences in annual forecast errors across a year's four fiscal quarters (DLS; Kang et al. 1994), we conduct separate analyses of forecast errors by quarter.[7] Table 1, Panel C presents forecast error distributional statistics by quarter. Consistent with prior research, the average forecast error is optimistic (DLS; FP; O'Brien 1988; Lys and Sohn 1990), and declines as the year progresses (Richardson et al. 2001). This decline is expected

---

[4] VL forecasted and actual EPS values exclude, "non-recurring or one-time gains and losses, which are noted in the footnotes." (See Value Line 2001.)

[5] This technique effectively adjusts EPS forecasts for stock splits, stock dividends, and sales of new and treasury shares.

[6] This value approximates beginning-of-the-year market value of equity and is available from the VL database. Similar results are obtained when we analyze the data scaled by book value.

[7] At the suggestion of a reviewer we also considered within-quarter forecast horizons by including an additional independent variable measuring time from the forecast date to quarter end. In none of our analyses did this additional variable lead to material changes in coefficient estimates, observed significance levels, or inferences regarding the association between forecast error and earnings predictability.

TABLE 1
Distribution Statistics

Panel A: Sample Forecasts by Year[a]

Year	Frequency	Percent
1987	1,639	5.57
1988	1,798	6.11
1989	1,840	6.25
1990	2,048	6.96
1991	2,377	8.08
1992	2,064	7.01
1993	2,206	7.50
1994	1,731	5.88
1995	2,814	9.56
1996	2,832	9.62
1997	2,655	9.02
1998	2,610	8.87
1999	2,818	9.57
Total	29,432	100.0

Panel B: Industry Breakdown[a]

SIC Code	Industry	Observations	Percentage
10	Metal Mining	733	2.5
12	Coal Mining	519	1.8
20	Food and Kindred Products	809	2.7
26	Paper	995	3.4
27	Printing	869	3.0
28	Plastics, Pharmaceuticals, Chemical	2,514	8.5
29	Petroleum Refining, Paving, Roofing	731	2.5
33	Iron and Steel	436	1.5
35	Machinery and Equipment	2,403	8.2
36	Electrical and Electronics	1,441	4.9
37	Transportation Equipment	1,202	4.1
38	Measuring Instruments	500	1.7
48	Communications	861	2.9
49	Utilities	3,594	12.2
60	Depository Institutions	1,799	6.1
61	Nondepository Credit Institutions	786	2.7
63	Insurance	1,516	5.2
73	Business Services	615	2.1
80	Health Care	1,081	3.7
	Other	6,028	20.5
	Total	29,432	100.0

(continued on next page)

TABLE 1 (continued)

Panel C: Annual Forecast Error as a Percentage of Market Value by Forecast Quarter[b]

Quarter	Observations	Mean	Median	Std. Dev.
First	7,213	−0.008	−0.002	0.035
Second	7,268	−0.007	−0.001	0.031
Third	7,556	−0.005	−0.001	0.026
Fourth	7,395	−0.003	−0.000	0.019
Total	29,432			

Panel D: Distributional Statistics for Independent Variables: 29,432 Firm-Quarter Observations[c]

	Mean	Median	Std. Dev.	Min	Max
UNPRED	713.661	695.000	438.597	1.000	1551.000
EARNINGS	0.062	0.065	0.056	−0.475	0.480
SIZE	3.151	3.132	0.668	1.126	5.516
TIMELY	3.031	3.000	0.9228	1.000	5.000

[a] All values obtained from *VL Estimates and Projections File.*

[b] Forecast error defined as actual annual earnings less forecasted earnings scaled by market value at the first quarter forecast date. All values obtained from *VL Estimates and Projections File.* All mean and median values are significant at the 0.001 level.

[c] All variables relate to firm-specific values over the period 1987 to 1999.

UNPRED = Value Line predictability ranking at the forecast date = ValuGauge Relative Position measure (1 to 1,600 with 1 being most predictable). For the years 1987 to 1999 the highest predictability ranking issued was 1,580, but due to the elimination of extreme forecast error and earnings values our sample maximum is less;

EARNINGS = annual earnings scaled by market value of common equity at the first fiscal quarter's forecast date;

SIZE = common logarithm of the market value of equity, market value of equity = beginning of the year outstanding shares × Value Line reported price per share on the first fiscal quarter's forecast date; and

TIMELY = Value Line timeliness ranking at the forecast date. The Value Line timeliness ranking ranges from 1 to 5, with a rank of 1 being the most favorable.

because early quarters' earnings are available when late quarter annual earnings forecasts are issued.

III. METHOD

Model

To examine the association between forecast error and earnings predictability we model annual forecast error as a function of earnings predictability, earnings level, firm size, and Value Line Timeliness Rank. Omitting firm and year subscripts for brevity, we obtain:

$$FE = \beta_0 + \beta_1 \times UNPRED + \beta_2 \times EARNINGS + \beta_3 \times SIZE + \beta_4 \times TIMELY + \varepsilon.$$

In the remainder of this section we define the variables and describe their measurement.[8]

---

[8] Our model is similar to DLS's with two notable exceptions. First, we include *EARNINGS* as an explanatory variable, while DLS do not. Second, while DLS include I/B/E/S analyst following (*FOLLOWING*) and the cross-products *FOLLOWING* × *UNPRED* and *SIZE* × *UNPRED*, we exclude these from our model because (1) *SIZE* and *FOLLOWING* are highly correlated, (2) the inclusion of these variables is atheoretical, (3) we obtain the same results whether we exclude or include these terms in the model, and (4) exclusion simplifies presentation of the results.

Table 1, Panel D, reports distributional statistics for the model's independent variables. We measure timeliness ranking, unpredictability, and forecast error on a quarterly basis to reflect the transitory nature of these variables. We measure firm size and actual earnings on an annual basis because firm size is relatively stable and we focus on errors in forecasting annual earnings. We define forecast error, *FE*, as actual annual earnings minus forecasted annual earnings, and scale forecast error by the market value of equity.

Earnings Predictability (UNPRED)

Value Line assesses the predictability of earnings based on the standard deviation of the percentage change in quarterly earnings from one year to the next, over the preceding eight years, with the most recent years weighted more heavily than earlier years. Value Line makes special proprietary adjustments for comparisons around zero earnings and when the sign of earnings changes (Value Line 2000, 2001). Value Line ranks the computed measures from 1 to nearly 1,600, where the most unstable earnings stream receives a value near 1,600 and the most stable earnings stream receives a value of 1. These ranks serve as the basis for VL's two published measures of earnings predictability. The first measure, the ValuGauge Relative Position, is the rank from 1 to nearly 1,600 and is available in the *VL Estimates* file. The second measure, the Value Line Predictability Index, further groups the initial ranking into 20 ranked groups valued from 5 to 100, in increments of 5, where 100 represents the highest level of predictability and 5 represents the least predictable earnings. We use the more precise ValuGauge Relative Position as our measure of earnings unpredictability (*UNPRED*), but note that use of the Value Line Predictability Index yields identical inferences. We use the VL measures of earnings predictability for two reasons. First, DLS show the VL Predictability Index has greater explanatory power in their model than the alternative earnings predictability measures they derive from time-series models, and second Luttman and Silhan (1993, 1995) show the Value Line Predictability Index is inversely related to subsequent absolute VL forecast accuracy, confirming that earnings are more difficult to forecast for low predictability firms.[9]

Earnings Level (EARNINGS)

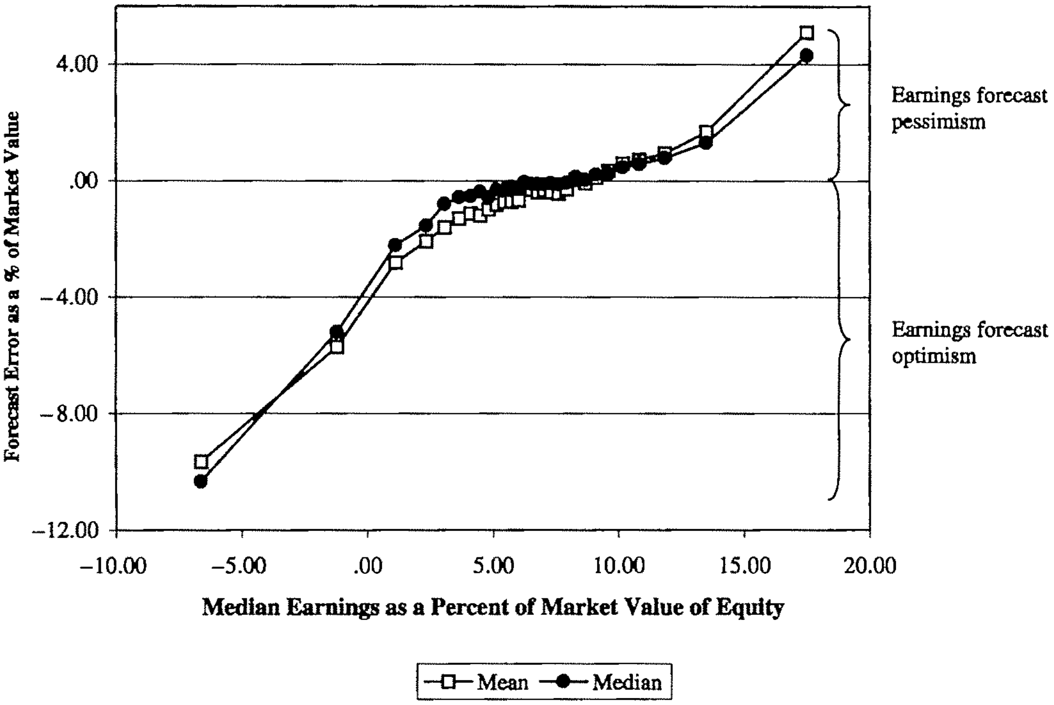
We measure *EARNINGS* as realized earnings for the year scaled by equity market value at the first fiscal quarter's forecast date. Accumulating evidence shows that forecast error varies with the level of earnings (e.g.,. Butler and Saraoglu 1999; Brown 2001) and several studies find increasing forecast optimism (pessimism) as the level of firm earnings declines below (increases above) the average earnings for all firms (Hwang et al. 1996; Eames et al. 2002; Eames and Glover 2002). This association between forecast error and the level of actual earnings is not surprising since forecast error equals actual earnings less forecasted earnings. Previous studies attribute this association to earnings management (Abarbanell and Lehavy 2002), as well as earnings shocks and a tendency of analysts' forecasts to exhibit excess regression to the mean (Eames et al. 2001).

Figures 1–3 illustrate the paired empirical relationships among earnings level, VL forecast error, and earnings predictability for first quarter forecast observations, respectively. We obtain similar relationships for the other quarters (not reported). For Figure 1, we sort

---

[9] DLS suggest that a limitation of the VL predictability measures is that the indices vary over time for a specific firm. We consider this variability an advantage because it is unlikely that firm-specific earnings predictability remains constant over multiple year periods, as DLS assume in their study. DLS (1998, 289) further warn that the VL predictability measure is "potentially subject to strategic behavior." However, this concern is unfounded as there is no strategic component of the measure because VL uses a purely mechanical computing process.

FIGURE 1
Value Line Forecast Error ($E - F$) by Earnings Level



Mean and median scaled Value Line forecast error (actual annual earnings less forecasted annual earnings, for forecasts issued in the first fiscal quarter) by median scaled annual earnings for 29 groups based on ranks of annual earnings. The highest earnings group has 213 observations and all others have 250 observations. We scale earnings and forecast values by market value of equity at the first fiscal quarter's forecast date. Similar results are obtained for quarters 2–4 (not depicted) and when we use beginning of the year book value as the scaler.

first quarter firm-year observations into portfolios based on earnings level, with 250 observations in all portfolios except the highest earnings portfolio. Then, for each portfolio, we plot the mean and median error in first quarter forecasts of annual earnings on the portfolio's median annual earnings. The figure shows that VL analysts' forecasts exhibit increasing optimism for scaled earnings below the median earnings value (6.5 percent) and increasing pessimism for earnings above the median.[10] The observed association is similar to that observed by Eames et al. (2002) for Zacks broker-analyst data and confirms the association for a population of VL analysts not subject to business generating incentives. The mean and median forecast optimism observed in Table 1, Panel C is driven by the fact that forecast optimism at lower earnings levels dominates the pessimism at higher earnings levels. Much of the previous forecast error research regarding the management relations hypothesis focuses on the overall mean optimism in forecast error. However, the pessimism illustrated in Figure 1 is not predicted or explicable by the management relations hypothesis.

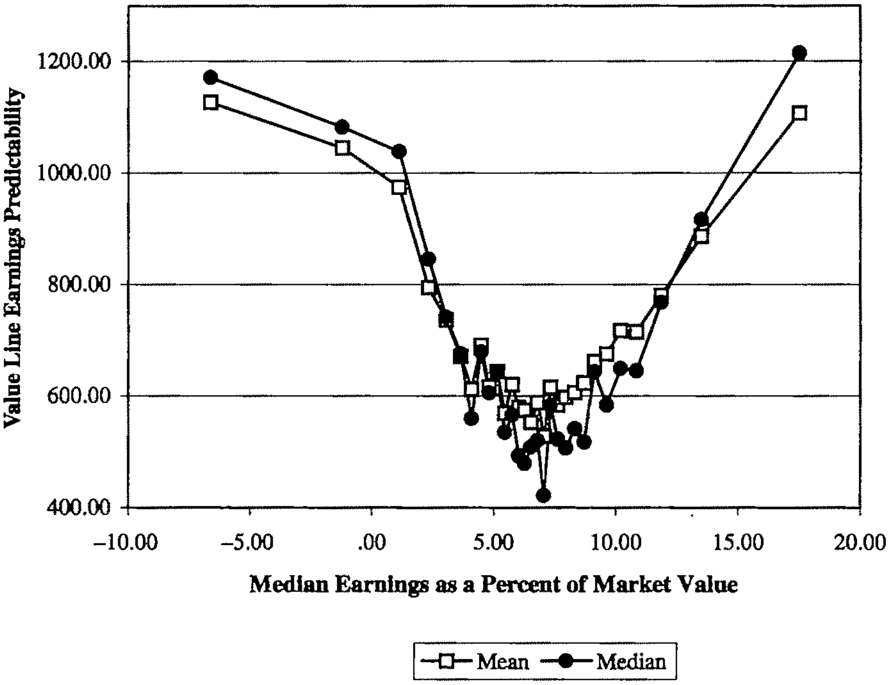
---

[10] Since $FE_{it} = E_{it} - F_{it}$, it appears there is an obvious algebraic association between FE_{it} and E_{it}. However, because F_{it} must be viewed as a function of E_{it} we do not necessarily expect a positive association between forecast error and earnings level, and such an association only exists under restrictive assumptions regarding the behavior of F_{it} with respect to E_{it}. To see this, we set $FE_{it} = E_{it} - F_{it} = E_{it} - (B_0' + B_1'E_{it} + e_{it}) = B_0 + B_1E_{it} + e_{it}$, differentiate with respect to E_{it}, obtain $1 - (\partial F/\partial E) = B_1$, and note that B_1 is positive only if $\partial F/\partial E < 1$, and $B_1 = 1$ only if $\partial F/\partial E = 0$. There is no basis for asserting either of these conditions *a priori*. For our sample $\partial F/\partial E = 0.57$ and is significantly different from zero at the 1 percent level. Thus, the association between forecast error and earnings is not algebraic, but rather, a systematic relationship between forecast errors and earnings requires systematic analyst behavior, such as an inability to forecast some component or portion of earnings.

While we document a significant association between earnings and forecast error, it is important to control for earnings level only if it is also associated with earnings predictability. In Figure 2, we plot mean and median *UNPRED* on portfolio median scaled earnings. The V-shape in Figure 2 shows that extreme earnings are least predictable and earnings near the median are most predictable. A joint consideration of Figures 1 and 2 indicates the potential for differing associations between forecast error and earnings predictability above and below median scaled earnings of 6.5 percent.

In Figure 3 we create separate plots for scaled earnings observations above (Panel A) and below (Panel B) 6.5 percent, the median value of earnings (see Table 1, Panel D).[11]

FIGURE 2
Value Line Earnings Predictability by Earnings Level



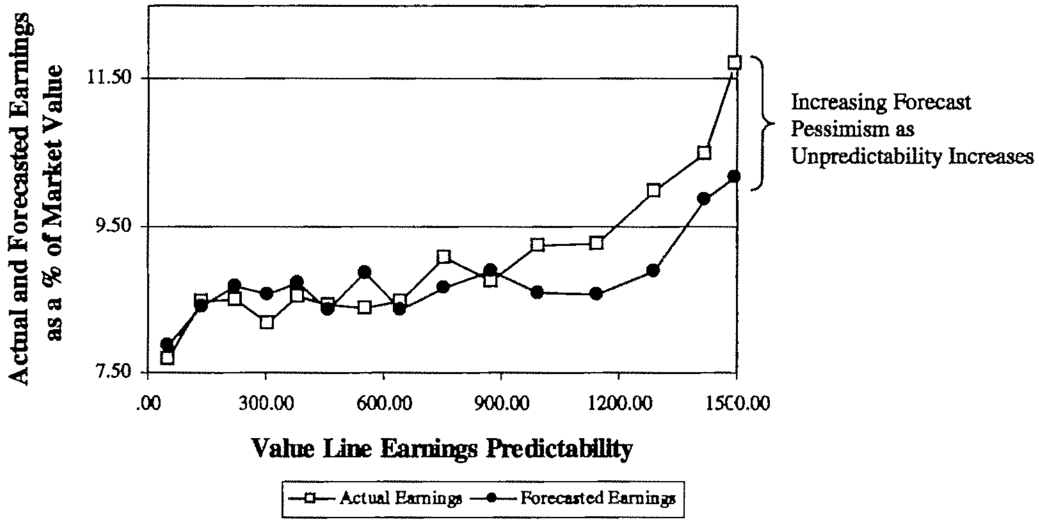
Mean and median Value Line earnings predictability (with smaller values indicating greater predictability) by median scaled annual earnings for 29 portfolios based on ranks of annual earnings and first fiscal quarter observations of earnings predictability. The highest earnings portfolio has 213 observations and all others have 250 observations. We scale earnings by the market value of common equity at the first fiscal quarter's forecast date. Similar results are obtained for quarters 2–4 (not depicted) and when beginning of the year book value is used as the scaler.

---

[11] Inferences from the figures and subsequent analyses are identical for a range of earnings level breaks from at least 6 to 7 percent.

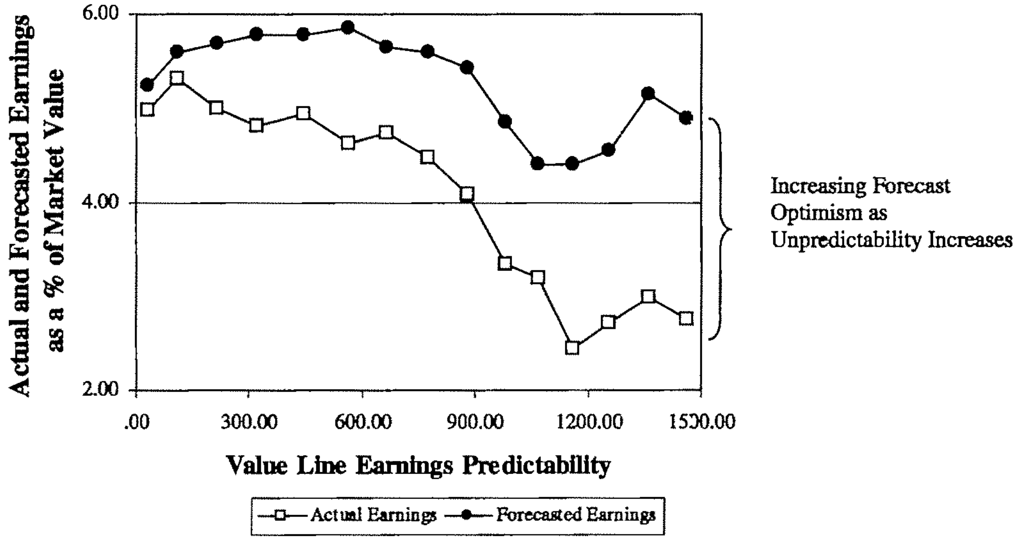
FIGURE 3
Actual and Value Line Forecasted Earnings Level by Value Line Earnings Predictability

Panel A: Earnings Above 6.5% (Median Value) of Market Value of Common Equity



Increasing Forecast Pessimism as Unpredictability Increases

Panel B: Earnings Below 6.5% (Median Value) of Market Value of Common Equity



Increasing Forecast Optimism as Unpredictability Increases

Median scaled actual annual earnings and first quarter forecasts of annual earnings by Value Line earnings predictability (with smaller values indicating greater predictability) for groups based on whether earnings exceed median value and earnings predictability rank. Panel A (Panel B) reflects 3,603 (3,610) observations grouped into 15 (15) equal-sized categories. We scale actual and forecasted earnings by market value of common equity at the first fiscal quarter's forecast date. Similar results are obtained for quarters 2–4 (not depicted) and when beginning of the year book value is used as the scaler.

For scaled earnings above 6.5 percent (Panel A), we find increasing forecast pessimism for increasingly unpredictable earnings. For earnings values below 6.5 percent (Panel B), we find increasing forecast optimism as earnings becomes more unpredictable. Figure 3 suggests that the basic association between VL earnings predictability and analyst forecast errors reverses above and below median earnings. While the pattern for earnings below 6.5 percent (Panel B) is consistent with the DLS theory, the results for earnings above 6.5 percent are not. Thus, Figure 3, Panel A suggests that DLS's results do not generalize to approximately half of the firms in our sample. We find the argument that analysts would attempt to curry favor with increasingly optimistic forecasts for increasingly unpredictable earnings *only* when managers report earnings below the median counter-intuitive and suggestive of model misspecification in the DLS study.

Firm Size (SIZE)

DLS control for firm size because it can proxy for the amount of public and private information available for a firm (Atiase 1987), which implies smaller forecast errors and greater earnings predictability for larger firms. We control for firm size measured as the common logarithm of market value, computed as the product of VL reported price per share on the first quarter forecast date and the number of shares outstanding at the end of the preceding fiscal year.

Timeliness Ranking (TIMELY)

To control for the previously identified positive association between forecast optimism and the Value Line Timeliness Rank and for consistency with DLS, we include the timeliness rank, *TIMELY*, in our model. The VL Timeliness Rank is VL's estimate of the firm's relative firm stock price performance over the next six to 12 months, on an integral scale of from 1 (highest) to 5 (lowest) (Value Line 2001). Rather than relying on analyst input, VL generates timeliness ranks via a proprietary computational algorithm, which ranks price performance relative to the other stocks VL follows.

Analyses of Five-Year Average Values and Individual Forecast Observations

DLS base their analysis on five-year average values for the period 1989 to 1993, to facilitate their estimation of earnings predictability and to reduce the likelihood of overstating significance as a result of serial correlation in the error term (DLS 1998, footnote 2). This averaging approach reduces the likelihood of overstating significance, but at the cost of losing information in the averaging process. The use of five-year average values assumes the variables are reasonably constant over the study period and that average values are interpretable in the context of the research question. However, for the 282 first-quarter observations for which values are available in all years from 1989 to 1993, not a single firm exhibits constant forecast error or predictability, while only 15 firms exhibit constant timeliness. Furthermore, 50 percent of firms exhibit firm-specific standard deviations for *UNPRED*, *FE*, and *TIMELY* in excess of 79, 1.4 percent, and .70, respectively, suggesting substantial intrafirm temporal variability in the data.[12] Moreover, it is difficult to interpret average values in our research context. An average timeliness rank of 3 may reflect a series of timeliness ranks never containing a 3 or all 3s. Similarly, an average forecast error of zero may result from a mix of large positive and negative individual errors.

---

[12] Affleck-Graves and Mendenhall (1990) report that, for the period 1982 to 1986, 51.5 percent of followed firms maintained their timeliness rank for a period of 13 weeks or less, only 8 percent maintained their rank for at least a year, and in an average week 5.3 percent of followed firms changed their rank. For a sample of 5 weeks covering September and October 1998 we find that in the average week 3.2 percent of firms experience a change in timeliness rank.

For consistency with DLS's method and study period, we first present results based on five-year average values over the period 1989 to 1993. However, considering the difficulty of interpreting results based on five-year averages, our primary analyses focus on individual forecasts of annual earnings, considered separately by forecast quarter. Our primary analyses are based on forecast observations over the period 1987 to 1999. We expect to find *no* association between forecast error and earnings predictability after controlling for the level of earnings. Consequently, serial correlation in the errors in our analysis of individual earnings forecasts may bias against our prediction of no relation.[13]

IV. RESULTS

Analyses with Five-Year Average Observations

For analysis of the five-year average values, like DLS, we limit the sample to firms for which VL forecasts are available in all quarters from 1989 to 1993, and do not eliminate or truncate outliers. This yields a sample of 255 observations for each quarterly forecast horizon.[14] The distribution of industries across this sample mimics that of DLS, for example, utilities represent 19 percent of the DLS sample and 17 percent of our sample.[15] The distribution of firm-specific average forecast errors, however, differs substantially between the samples. Mean forecast errors for our sample vary from −0.012 in the first quarter to −0.001 in the fourth quarter, while similar values for the DLS sample are −0.030 and −0.015, respectively.[16] Furthermore, our forecast errors exhibit much less range than the DLS sample. For each quarter our maximum forecast error was 1.40 and minimums ranged from −0.26 in the fourth quarter to −0.39 in the third quarter. In contrast, DLS's maximum was 1.43, and minimum values range from −2.7 in the first quarter to −2.4 in the third and fourth quarters. Thus, the DLS sample includes extreme forecast errors that were 2.7 times the magnitude of market value.[17]

Table 2 displays the regression results for five-year average values, both with and without the *EARNINGS* control variable. Results for the model excluding *EARNINGS* are generally consistent with DLS's results (DLS 1998, Table 4). For the first two quarters,

---

[13] We conduct sensitivity analyses to address the potential impact of both time-series and cross-sectional correlation in the error term and find no evidence of a significant association between *FE* and *UNPRED*. To limit the effect of serially correlated errors, we individually estimate the model for each of the 52 quarters from 1987 to 1999. In Figure 3, and in analyses reported later, we demonstrate the importance of partitioning on *EARNINGS* and including *EARNINGS* in the model; thus, for the sake of brevity, we limit our discussion to analyses incorporating these controls. Among the resulting 104 (2 × 52 quarters) regressions, 79 yield coefficients on the *UNPRED* that are nonsignificant at the 5 percent level. Of the remaining, 25 coefficients, 9 are significantly positive and 15 are significantly negative at the 5 percent level. We also estimate the model across all firms and years while treating year and firm-specific effects first as fixed effects then second as random components (Kennedy 1992). Generally, we find no significant or consistent association between forecast error and *UNPRED*. Applying the random-effects model to earnings observations below 6.5 percent we find evidence of significant pessimism as earnings predictability decreases. This result is not consistent with DLS or the nonsignificant results in Table 3, Panel B.

[14] The DLS sample consists of 239 firm-specific five-year average observations for each quarter. Requiring that observations for all variables be available for all quarters of the 20-quarter sample period significantly limits the sample. A wide variety of factors, such as the temporary suspension of timeliness ranks, nonreporting of earnings predictability, changes in the firms covered, and lack of I/B/E/S International Inc. reporting in a single year significantly reduce the sample size. Without such stringent requirements, we would have obtained a sample of 820 five-year average observations.

[15] This 17 percent is higher than the 12.2 percent for utilities in our sample of firm-year observations over the period 1987 to 1999 (Table 1, Panel B) due to the substantial continuity in data required for the five-year sample and the relative stability of firms in the utilities industry.

[16] The mean forecast error reported in DLS is significantly higher than that reported in related research (e.g., Brown 2001; FP; Eames et al. 2002; Eames and Glover 2002).

[17] While we gather VL data from the same years as DLS, the extreme observations reported by DLS may have been corrected or dropped by VL.

TABLE 2
Association between Five-Year Average Earnings Predictability and Five-Year Average Forecast Error Using Firm-Quarter Forecasts of Annual Earnings (1989–1993)

$$FE = \beta_0 + \beta_1 \times UNPRED + \beta_2 \times EARNINGS + \beta_3 \times SIZE + \beta_4 \times TIMELY + \varepsilon$$

	Without *EARNINGS* Control				With *EARNINGS* Control			
Independent Variable	**1st Qtr**	**2nd Qtr**	**3rd Qtr**	**4th Qtr**	**1st Qtr**	**2nd Qtr**	**3rd Qtr**	**4th Qtr**
Intercept	0.055	0.077	0.044	0.054	−0.047***	−0.046***	−0.062***	−0.065***
UNPRED (× 10^4)	−0.435***	−0.383***	−0.350**	−0.252*	0.095**	0.133**	0.173***	0.250***
SIZE	−0.001	−0.004	−0.004	−0.006	0.000	−0.003	−0.002	−0.004
TIMELY	−0.012	−0.016	−0.005	−0.007	−0.008**	−0.005	−0.000	0.003
EARNINGS					0.872***	0.842***	0.846***	0.804***
Adjusted R^2 (%)	2.8	2.4	1.2	0.29	90.7	88.0	87.7	82.8
F-statistic	3.39**	3.08**	2.06	1.24	621.83***	468.15***	452.53***	306.28***
Observations	255 for each quarter				255 for each quarter			

---

*, **, *** Significance at the <0.10, <0.05, and <0.01 levels, respectively.

All variables relate to firm-specific five-year average values over the period 1989 to 1993.

FE = average forecast error, forecast error = annual earnings less forecast earnings, all scaled by market value of common equity at the first fiscal quarter's forecast date;

UNPRED = average earnings unpredictability, Value Line (VL) ValuGauge Relative Position measure of earnings predictability. This measure ranges from 1 to 1,600, where 1 denotes highly predictable earnings;

EARNINGS = average earnings scaled by market value of equity at the first fiscal quarter's forecast date;

SIZE = common logarithm of the average market value of equity, market value of equity = beginning of the year outstanding shares multiplied by VL reported price per share on the first fiscal quarter's forecast date; and

TIMELY = average VL timeliness ranking. The VL timeliness ranking ranges from 1 to 5, with a rank of 1 being the most favorable.

coefficient estimates for *UNPRED* are negative and significant at the 1 percent level. For the third quarter the coefficient is again negative, but significant at the 5 percent level. For the fourth quarter the coefficient is negative, but only modestly significant (10 percent level). We obtain less significance and lower R^2 values than DLS, and attribute these differences to the extreme observations in the DLS sample.

Including *EARNINGS* in the regression model leads to very different conclusions. First, the coefficient estimates for *EARNINGS* are uniformly positive and highly significant, consistent with prior research (e.g., Eames et al. 2002; Eames and Glover 2002). With *EARNINGS* in the model, the explanatory power increases substantially and all reported R^2 values exceed 82 percent. We now find a positive and significant coefficient on *UNPRED* for all quarters; a pattern that is inconsistent with DLS's prediction that as earnings predictability decreases forecast optimism increases.

Analysis of Individual Firm-Quarter Forecasts of Annual Earnings

Considering the serious limitations of analyses of firm-specific five-year average values, our primary analyses are based on firm-quarter forecasts of annual earnings. Reviewing Figures 1 to 3, it appears that the linear model and *EARNINGS* control considered in the five-year average analysis is not appropriate. Considering the switch in the sign of the association between forecast error and earnings predictability above and below the median earnings, we construct two subsamples by partitioning our data at the median level of *EARNINGS*, and separately analyze each subsample.[18]

Table 3, Panel A reports regression results for observations with scaled earnings greater than 6.5 percent. Excluding *EARNINGS* from the model, the coefficient on *UNPRED* is positive and significant for each quarter, indicating as unpredictability increases, forecast pessimism increases—just the opposite of DLS's hypothesized behavior. Adding *EARNINGS* to the model, the coefficients on *UNPRED* are inconsistent in sign and not significant in the final three quarters, while in the first quarter the coefficient is positive and significant, again inconsistent with the DLS hypothesis. Thus, for earnings levels above the median we find no evidence supportive of the DLS hypothesis.

Table 3, Panel B reports regression results for observations with scaled earnings less than 6.5 percent. Without *EARNINGS* in the model the coefficients on *UNPRED* are negative and highly significant for each quarter, indicating as unpredictability increases, forecast optimism increases. This pattern is consistent with DLS's hypothesized behavior. However, with *EARNINGS* in the model the coefficients on *UNPRED* are not significant in any of the last three quarters, and only modestly significant (10 percent level) in the first quarter. Thus, even when we focus on the portion of the earnings distribution seemingly most consistent with the DLS hypothesis, we find that prior evidence in DLS supporting the theory that analysts issue intentionally optimistic forecasts as earnings unpredictability increases is likely attributable to a failure to adequately consider underlying relationships among earnings level, forecast error, and earnings predictability.

V. SUMMARY AND CONCLUSIONS

Relying on the management relations hypothesis developed by Francis and Philbrick (1993), Das et al. (1998) suggest that as earnings become less predictable, analysts have

---

[18] The dichotomous control is in the same spirit as Brown (2001) who considers a temporal analysis of earnings surprises and finds decidedly different results for profit and loss firms. It is readily apparent from an inspection of both Figures 1 and 2 that the direction of forecast error switches around median and mean earnings. Thus, when analyzing forecast error the appropriate "switch point" is median/mean earnings and not profit and loss.

TABLE 3
Association between Earnings Predictability and Analysts' Forecast Error

$$FE = \beta_0 + \beta_1 \times UNPRED + \beta_2 \times EARNINGS + \beta_3 \times SIZE + \beta_4 \times TIMELY + \varepsilon$$

Panel A: Using Firm-Quarter Forecasts of Annual Earnings with Scaled Earnings Greater than 6.5% (1987–1999)

	Without *EARNINGS* Control				With *EARNINGS* Control			
Independent Variable	**1st Qtr**	**2nd Qtr**	**3rd Qtr**	**4th Qtr**	**1st Qtr**	**2nd Qtr**	**3rd Qtr**	**4th Qtr**
Intercept	0.020***	0.009***	0.007***	0.006***	0.021***	0.024***	−0.014***	−0.004**
UNPRED (× 10^4)	0.106***	0.079***	0.040***	0.025***	0.024***	0.004	−0.004	0.005
SIZE	−0.000	−0.001	−0.000	−0.001***	0.001**	0.000	0.000	−0.009**
TIMELY	−0.007***	−0.003***	−0.002***	−0.001***	−0.006***	−0.002***	−0.001***	−0.001**
EARNINGS					0.392***	0.330***	0.204***	0.093***
Adjusted R^2 (%)	10.6	4.7	2.9	1.7	30.1	25.8	16.3	6.2
F-statistic	142.73***	60.01***	37.64***	21.63***	389.09***	316.65***	183.09***	60.52***
Observation	3,603	3,629	3,738	3,621	3,603	3,629	3,738	3,621

(*continued on next page*)

TABLE 3 (continued)

Panel B: Using Firm-Quarter Forecasts of Annual Earnings with Scaled Earnings Less than 6.5% (1987–1999)

	Without *EARNINGS* Control				With *EARNINGS* Control			
Independent Variable	**1st Qtr**	**2nd Qtr**	**3rd Qtr**	**4th Qtr**	**1st Qtr**	**2nd Qtr**	**3rd Qtr**	**4th Qtr**
Intercept	−0.019***	−0.017***	−0.013***	−0.006**	−0.037***	−0.039***	−0.027***	−0.014***
UNPRED (× 10^4)	−0.154***	−0.125***	−0.096***	−0.058***	−0.023*	0.015	0.005	0.009
SIZE	0.012***	0.010***	0.007***	0.004***	0.009***	0.007***	0.005***	0.003***
TIMELY	−0.009***	−0.007***	−0.005***	−0.003***	−0.007***	−0.004***	−0.003***	−0.002***
EARNINGS					0.436***	0.400***	0.265***	0.146***
Adjusted R^2 (%)	16.4	13.1	9.2	5.9	39.4	39.1	28.2	19.0
F-statistic	236.32***	183.45***	130.30***	79.4***	586.63***	583.96***	376.39***	222.17***
Observation	3,610	3,639	3,818	3,774	3,610	3,639	3,818	3,774

*, **, *** Significance at the <0.10, <0.05, and <0.01 levels, respectively.

All variables relate firm-year observations for the period 1987–1999.

FE = forecast error, forecast error = annual earnings less forecast earnings, all scaled by the market value of common equity at the first fiscal quarter's forecast date;

UNPRED = earnings unpredictability at the forecast date, VL ValuGauge Relative Position measure of earnings predictability. This measure ranges from 1 to 1,600, where 1 denotes highly predictable earnings;

EARNINGS = earnings scaled by market value of equity at the first fiscal quarter's forecast date;

SIZE = common logarithm of the market value of equity, market value of common equity = beginning of the year outstanding shares * Value Line reported price per share on the first fiscal quarter's forecast date; and

TIMELY = Value Line timeliness ranking at the forecast date. The Value Line timeliness ranking ranges from 1 to 5, with a rank of 1 being the most favorable.

an incentive to issue optimistic forecasts to curry favor with management so the analyst can obtain management's private information. However, recent articles in the academic and popular press suggest that issuing intentionally optimistic earnings forecasts is not an effective means of currying favor with management because optimistic earnings forecasts lead to negative earnings surprises, which are often associated with negative market reactions.

We re-examine the association between Value Line forecast error and earnings predictability for the period 1987 to 1999 using a more extensive data set than Das et al. (1998). We are able to replicate Das et al.'s (1998) results with our sample using similar techniques and for the same sample period. However, our results indicate that earnings is an important control variable when investigating the association between forecast error and earnings predictability. We document an association between earnings level and earnings predictability where firms with the most predictable earnings typically report earnings near the sample median earnings level. Plots of the association between forecast error and earnings predictability above and below median earnings suggest that Das et al.'s (1998) theory and results do not generalize to firms reporting earnings above the median. When we control for the level of earnings by including earnings level as an independent regression variable, we find no significant association between forecast error and earnings predictability. Our results suggest that Das et al.'s (1998) results are spurious and were caused by omitting earnings level from the analysis. We conclude that there is no significant evidence of an inverse association between Value Line forecast error and earnings predictability.

REFERENCES

Abarbanell, J., and R. Lehavy. 2002. Biased forecasts or biased earnings? The role of earnings management in explaining apparent optimism and inefficiency in analysts' earnings forecasts. *Journal of Accounting and Economics* (2002 Conference).

Affleck-Graves, J., L. Davis, and R.R. Mendenhall. 1990. Forecasts of earnings per share: Possible sources of analyst superiority and bias. *Contemporary Accounting Research* 6 (Spring): 501–517.

Atiase, R. 1987. Market implications of predisclosure information: Size and exchange effects. *Journal of Accounting Research* 25 (Spring): 168–176.

Bartov, E., D. Givoly, and C Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 22 (June): 173–204.

Brown, L. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39 (September): 221–241.

Burgstahler, D., and M. Eames. 2002. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. Working paper, University of Washington.

Butler, K., and H. Saraoglu. 1999. Improving analysts' negative earnings forecasts. *Financial Analysts' Journal* 55 (May/June): 48–56.

Das, S., C. B. Levine, and K. Sivaramakrishnan. 1998. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review* 73 (April): 277–294.

Drucker, J., and J. Sapsford. 2002. Analyze this: Unhappy firm bites back. *Wall Street Journal* (November 22): C1.

Eames, M., S. Glover, and K. Stice. 2001. The relation between earnings and earnings forecast error. Working Paper, Brigham Young University.

———, and ———. 2002. Value Line timeliness ranks and Value Line earnings forecast errors. Working Paper, Santa Clara University.

———, ———, and J. Kennedy. 2002. The association between trading recommendations and broker-analysts' earnings forecasts. *Journal of Accounting Research* 40 (March): 85–104.

Francis, J., and D. R. Philbrick. 1993. Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research* 31 (Autumn): 216–230.

Hwang, L., C. Jan, and S. Basu. 1996. Loss firms and analysts' earnings forecast errors. *The Journal of Financial Statement Analysis* 1 (Winter): 18–29.

Jorge, M., and W. Rees. 2000. Three triangulated studies of financial analysts in Spain. Working paper, University of Glasgow.

Kang, S., J. O'Brien, and K. Sivaramakrishnan. 1994. Analysts' interim earnings forecasts: Evidence of the forecasting process. *Journal of Accounting Research* 32 (Spring): 103–112.

Kasznik, R., and M. McNichols. 2002. Does meeting expectations matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research* 40 (June): 727–759.

Kennedy, P. 1992. *A Guide to Econometrics*. Third edition. Cambridge, MA: MIT Press.

Kim, C., and S. Lustgarten. 1998. Broker-analysts' trade-boosting incentive and their earnings forecast bias. Working paper, Queen's College of the City University of New York.

Kinney, W., D. Burgstahler, and R. Martin. 2002. Earnings surprise "materiality" as measured by stock returns. *Journal of Accounting Research* 40 (December): 1297–1329.

Lopez, T., and L. Rees. 2002. The effect of meeting analyst forecasts and systematic positive forecast errors on the information content of unexpected earnings. *Journal of Accounting, Auditing, and Finance* 17 (Spring): 155–184.

Luttman, S., and P. Silhan. 1993. An empirical analysis of the Value Line earnings predictability index. *Journal of Applied Business Research* 9 (4).

———, and ———. 1995. Identifying factors consistently related to Value Line earnings predictability. *The Financial Review* 30 (August): 445–468.

Lys, T., and S. Sohn. 1990. The association between revisions of analysts' earnings forecasts and security-price changes. *Journal of Accounting and Economics* 13 (December): 341–363.

Matsumoto, D. A. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (July): 483–514.

O'Brien, P. C. 1988. Analysts' forecasts as earnings expectation. *Journal of Accounting and Economics* 10: 53–83.

Richardson, S., S. Teoh, and P. Wysocki. 2001. The walkdown to beatable analyst forecasts: The roles of equity issuance and insider trading incentives. Working paper, University of Michigan.

Skinner, D., and R. Sloan. 2001. Earnings surprises, growth expectations and stock returns or don't let an earnings torpedo sink your portfolio. Working paper, University of Michigan.

Value Line. 2000. Glossary of investment terms. Available at: http://valueline.com/sup_glossa.html.

———. 2001. *How to Invest in Common Stocks: The Complete Guide to Using the Value Line Investment Survey*. New York, NY: Value Line Publishing Inc.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 725–758

Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard

Christopher D. Ittner
David F. Larcker
Marshall W. Meyer
University of Pennsylvania

ABSTRACT: This study examines how different types of performance measures were weighted in a subjective balanced scorecard bonus plan adopted by a major financial services firm. Drawing upon economic and psychological studies on performance evaluation and compensation criteria, we develop hypotheses regarding the weights placed on different types of measures. We find that the subjectivity in the scorecard plan allowed superiors to reduce the "balance" in bonus awards by placing most of the weight on financial measures, to incorporate factors other than the scorecard measures in performance evaluations, to change evaluation criteria from quarter to quarter, to ignore measures that were predictive of future financial performance, and to weight measures that were not predictive of desired results. This evidence suggests that psychology-based explanations may be equally or more relevant than economics-based explanations in explaining the firm's measurement practices. The high level of subjectivity in the balanced scorecard plan led many branch managers to complain about favoritism in bonus awards and uncertainty in the criteria being used to determine rewards. The system ultimately was abandoned in favor of a formulaic bonus plan based solely on revenues.

Keywords: *balanced scorecard; subjective performance measures; nonfinancial performance measurement.*

Data Availability: *All data are proprietary. Confidentiality agreements prevent the authors from distributing the data.*

I. INTRODUCTION

This study investigates the use of subjectivity in reward systems containing multiple performance measures. Bonuses based solely on profits and other financial accounting numbers have been criticized for encouraging managers to sacrifice long-run

---

We appreciate the comments of the reviewer and seminar participants at the University of Alberta, University of Cambridge, Citibank Behavioral Science Research Council Conference, Georgia State University, University of New South Wales, University of Saskatchewan, Stanford University Summer Accounting Research Camp, The University of Texas at Austin, The University of Texas at Dallas, and University of Waterloo. This research was funded by the Citibank Behavioral Science Research Council, whose support is gratefully acknowledged.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted September 2002
Accepted February 2003

performance to increase short-term financial results, and thereby maximize their bonuses. To overcome the short-run orientation of accounting-based reward systems, many firms are implementing compensation plans that supplement financial metrics with additional measures in order to assess performance dimensions that are not captured in short-term financial results. These additional measures can take a variety of forms, ranging from quantitative, nonfinancial metrics, such as employee and customer survey results, to qualitative assessments of performance by the manager's superior.

One critical implementation issue that arises when incorporating multiple performance measures in reward systems is determining the relative weights to place on the various measures when determining bonuses. One option is to use a formula that explicitly weights each measure. Potential difficulties with this option include determining the appropriate weights to place on each measure, the "game-playing" associated with any formula-based plan, the possibility that bonuses will be paid even when performance is "unbalanced" (i.e., overachievement on some objectives and underachievement on others), and the likelihood that all relevant dimensions of managerial performance are not captured by the selected performance measures (e.g., Holmstrom and Milgrom 1991; Baker et al. 1994; Kaplan and Norton 1996).

A second option is to introduce subjectivity into the bonus award process. This subjectivity can take the form of flexibility in weighting quantitative performance measures when computing a manager's bonus, the use of qualitative performance evaluations, and/or the discretion to adjust bonus awards based on factors other than the measures specified in the bonus contract. Some theoretical work indicates that greater subjectivity can improve incentive contracting because it allows the firm to exploit noncontractible information that might otherwise be ignored in formula-based contracts, and to mitigate distortions in managerial effort by "backing out" dysfunctional behavior induced by incomplete objective performance measures (Baker et al. 1994; Baiman and Rajan 1995). However, other research suggests that subjectivity in reward systems can lower managers' motivation by allowing evaluators to ignore certain types of performance measures that are included in the bonus plan, permitting bonus payout criteria to change each period, and introducing favoritism and bias into the reward system (e.g., Prendergast and Topel 1993). As a result, managers will be less able to distinguish what constitutes good performance, less likely to believe that rewards are contingent on performance, and less convinced that performance criteria are being applied consistently across the organization.

Drawing upon economic and psychological studies on the choice of performance measures for performance evaluation and compensation purposes, we develop exploratory hypotheses regarding the weights placed on different types of performance measures (e.g., financial versus nonfinancial, quantitative versus qualitative, and input versus outcome) in subjective bonus computations. We test these hypotheses using quantitative and qualitative data gathered during an extensive, multiyear field investigation of a balanced scorecard bonus plan in the U.S. retail banking operations of Global Financial Services (GFS), a leading international financial services provider.

In 1993, GFS replaced a profit-based bonus plan for branch managers with a *formula-based* system that rewarded multiple accounting and growth measures once customer satisfaction and operational audit hurdles were achieved. This system changed rapidly during the nine quarters it was in use as the bank sought to eliminate gaming and promote performance across a broader set of measures. The formula-based plan was replaced in the second quarter of 1995 by a "balanced scorecard" bonus plan containing six categories of financial and nonfinancial performance measures, some of which were based on qualitative evaluations by the managers' supervisors. Unlike the earlier formula-based plan, *subjective* weightings were used to aggregate the various scorecard measures when determining overall performance evaluations and bonuses.

Although Kaplan and Norton (2001) cite the GFS bonus plan as an example of a scorecard-based reward system that prevented managers from underperforming on any of the scorecard dimensions, we find that the use of subjectivity in weighting the scorecard measures allowed supervisors to ignore many performance measures, even though some of these measures were leading indicators of the bank's two strategic objectives (financial performance and growth in customers). Instead, short-term financial performance measures become the primary determinant of bonuses. In addition, a large proportion of branch managers' performance evaluations was based on factors *other than* the scorecard measures, even though discretion to consider other factors was not a component of the bonus plan. The move from the formula-based system to the more subjective scorecard led many branch managers to complain about favoritism in bonus awards and uncertainty in the criteria being used to determine rewards, and caused corporate executives and human resource managers to question the scorecard's use for compensation purposes. Ultimately, the company abandoned the scorecard plan at the end of 1998 in favor of a commission-style system based on revenues.

Our study makes four contributions to the performance evaluation and compensation literatures. First, we extend cross-sectional studies on the use of subjectivity and discretion in bonus plans (e.g., Murphy and Oyer 2001; Gibbs et al. 2002). Whereas these studies focus on the factors explaining *who* includes subjectivity in compensation contracts, our study emphasizes *how* subjectivity is incorporated into performance evaluations and bonus awards. Second, we provide further evidence on the influence of informativeness on performance measure weighting. Prior studies on the relative weights placed on financial and nonfinancial performance measures (e.g., Bushman et al. 1996; Ittner et al. 1997) generally include proxies for the noise in financial measures, but do not include direct measures of the informativeness of *nonfinancial* measures due to data constraints. The detailed data in our study allow us to develop stronger tests of economic theories that the relative weights placed on performance measures other than financial results are a function of their informativeness. Third, our research complements psychology-based experimental work on the importance placed on various types of performance measures by examining whether their experimental results hold in an actual performance-evaluation setting. Finally, although the ability to generalize our results is limited by the analysis of only a single firm, we provide one of the first detailed studies of scorecard-based compensation plans.[1] Despite survey evidence that a growing number of firms are using balanced scorecards for compensation purposes (Kaplan and Norton 2001, Chapter 10), relatively little is known about the implementation issues associated with scorecard-based reward systems.

The remainder of the paper is organized in five sections. Section II reviews related research on subjective performance evaluation and performance measure weighting, and develops our exploratory hypotheses. The research setting for our study is described in Section III. The statistical tests for our research hypotheses are presented in Section IV. In Section V we provide detailed qualitative analysis regarding the success of the balanced scorecard implementation. The final section summarizes our research results and discusses implications for future research.

---

[1] See Malina and Selto (2001) and Campbell et al. (2002) for field-based studies examining other uses of the balanced scorecard. Also see Banker et al. (2000) for field evidence on the implementation of a compensation plan incorporating nonfinancial measures.

II. LITERATURE REVIEW AND HYPOTHESES

Performance Measure Use in Subjective Performance Evaluation

Kaplan and Norton (1996, 10) argue that balanced scorecards should reflect four types of measures: (1) financial and nonfinancial; (2) external (financial and customer) and internal (critical business processes, innovation, and learning and growth); (3) inputs/drivers and outcomes/results; and (4) objective, easily quantifiable measures and more subjective, judgmental measures. Although Kaplan and Norton (1996, 2001) provide little guidance on how to combine or "balance" these disparate measures when evaluating managerial performance, they conjecture that the balanced scorecard renders subjective reward systems "easier and more defensible to administer...and also less susceptible to game playing." (See Kaplan and Norton 1996, 220.)

Analytical research on the use of subjectivity in performance evaluation and compensation focuses on the benefits of subjective bonus awards (e.g., Baker et al. 1994; Baiman and Rajan 1995), the drawbacks of subjective performance evaluations (e.g., Prendergast and Topel 1996; MacLeod 2001), and the factors influencing the relative weights placed on subjective versus objective performance measures in incentive contracts (e.g., Murphy and Oyer 2001). Most of these models do not examine how different types of performance measures or different forms of subjectivity (i.e., flexibility in assigning weights to measures, use of qualitative performance evaluations, and/or discretion to incorporate other performance criteria) should be incorporated into subjective bonus awards. An exception is Murphy and Oyer's (2001) model, which suggests that the relative weight on subjective measures will be higher in privately held companies, larger companies with more top managers, less autonomous business units, companies with substantial growth opportunities, and companies where accounting profits and shareholder returns are less highly correlated. Their cross-sectional empirical tests of executive bonuses provide mixed support for these hypotheses.

A related empirical study of automobile dealerships by Gibbs et al. (2002) finds that subjectivity (defined as the presence of any subjective bonus payout or as "discretionary bonus" as a percent of total compensation) is positively related to departmental interdependencies, financial losses, the manager's tenure, and the achievability of formula-based bonuses. While these two empirical studies provide insight into *who* uses subjectivity in compensation contracts, they provide little insight into *how* subjectivity is applied or performance is evaluated when multiple types of performance measures are incorporated into the bonus contract.[2]

Other studies address the relative importance placed on the various types of measures highlighted by Kaplan and Norton (1996). These studies fall into two research streams. The first stream focuses on economic models of incentive contracting. Economics-based agency models emphasize the role of performance measure choice in aligning agents' goals with those of the principal, and indicate that the choice of performance measures in incentive contracts should be a function of the informativeness (or incremental information content) of each measure regarding the worker's action choices (e.g., Holmstrom 1979; Banker and Datar 1989; Feltham and Xie 1994; Hemmer 1996; Lambert 2001).

A second research stream adopts a psychological perspective. These studies examine how human information-processing capabilities and decision strategies influence the types of information individuals use when assessing performance. These behavioral experiments suggest that issues such as information overload and cognitive biases can play a significant role in the relative weights placed on different types of balanced scorecard measures (e.g., Lipe and Salterio 2000, 2002). In particular, this research finds that evaluators frequently place greater or exclusive emphasis on certain types of measures, even when other types of measures also provide relevant information on the subordinate's performance.

---

[2] Murphy and Oyer's (2001) tests examine the use of individual performance appraisals and discretionary allocation of bonus awards, but do not examine the use of specific types of performance measures. Gibbs et al. (2002) use an aggregate measure of subjectivity, which their survey defines as bonuses awarded based on a supervisor's subjective judgment of a manager's performance.

Economics-Based Hypotheses

Informativeness

Economics-based agency models focus on the role of performance measures in promoting congruence between the principal's objective and that of the agent (e.g., Holmstrom 1979; Banker and Datar 1989; Lambert 2001). Two primary insights from these models are that compensation contracts should include any (costless) measure that carries incremental information on the agent's actions, and that the relative weight placed on an individual measure should be a function of the measure's signal-to-noise ratio, as reflected in its sensitivity (or the change in its mean value in response to a change in the agent's action) and precision (or the inverse of the variance in the measure given the agent's action).

Although many of these models say little about the *specific* types of performance measures that should be included in compensation contracts, several studies extend these papers to investigate the role of nonfinancial measures (Feltham and Xie 1994; Hauser et al. 1994; Hemmer 1996). These models suggest that financial measures alone are unlikely to be the most efficient means to motivate employees, and demonstrate how incentives based on nonfinancial measures can improve contracting by incorporating information on agents' actions that is not fully captured in contemporaneous financial results.

A number of cross-sectional empirical studies draw upon informativeness theories when examining the relative weights placed on individual, nonfinancial, or subjective performance measures (e.g., Bushman et al. 1996; Ittner et al. 1997; Murphy and Oyer 2001; Ittner and Larcker 2002). These studies typically use two approaches to test the models' predictions. First, the effects of noise on performance measure choice are examined using the variance in objective, financial measures. In these tests, the weight placed on financial measures is predicted to decrease in their noisiness, while the weight on other types of measures is predicted to increase. Second, proxies for factors that are expected to influence the informativeness of individual, nonfinancial, or subjective measures (e.g., growth opportunities, strategy, product life cycle, etc.) are examined, with the weight placed on these measures expected to be higher when the informativeness proxies are greater.

The results from these studies are mixed. While proxies for the factors predicted to influence informativeness are generally associated with increased weights on these measures, proxies for the noise in financial measures tend to have little association with measurement choices. However, a significant limitation of these analyses is the lack of data on nonfinancial or subjective performance dimensions, forcing researchers to use indirect proxies for the measures' informativeness.

Consistent with the nonfinancial performance measure and balanced scorecard literatures, we assume that current financial measures are potentially incomplete, and that other indicators of future financial performance can provide incremental information on the manager's actions (e.g., Feltham and Xie 1994; Hemmer 1996; Kaplan and Norton 1996, 2001). We also assume that measures that are more predictive of future performance provide greater information on the congruence between the agent's actions and the outcomes desired by the principal. In an agency setting, the coefficients (or weights) associated with the nonfinancial performance measures in the structural model linking nonfinancial measures to future financial results (i.e., the "business model") and the coefficients (or weights) used in the agent's compensation contract will be identical if the agent is *risk neutral* (e.g., Datar et al. 2001). However, theoretical models by Gjesdal (1981) and Datar et al. (2001) also indicate that when the agent is *risk averse*, the coefficients in the business model will not be identical to the coefficients in the compensation model, complicating our hypothesis development and subsequent empirical tests.

Although the latter result is a potential concern, additional factors suggest that the statistical associations between nonfinancial measures and subsequent financial results are likely to provide reasonable proxies for the nonfinancial measures' informativeness in our setting. First, agency theory suggests that it would be unusual to observe a nonfinancial performance measure with a substantial positive coefficient in the business model (i.e., a measure that is highly predictive of future improvements in financial performance), but a zero or negative coefficient in the compensation model. This would only occur if the nonfinancial performance measure "corrects" the measurement error in another performance variable included in the model (i.e., a type of relative performance measurement), or if the use of the nonfinancial measure imposes too much risk on a risk-averse agent. We do not believe that our research setting exhibits either of these attributes.

Second, the lack of direct correspondence between the weights in the business model and the compensation model only arises in models that do not allow for private information by the agent. Once private information is incorporated into the formal agency model, the coefficients in the business model become quite similar (but not identical) to the coefficients in the compensation function. Since branch managers are likely to have private information about their operations and marketplace that is not available to their superiors, the comparison of weights between the business and compensation models should provide a reasonable test of informativeness theories. More specifically, if informativeness is a significant determinant of performance measurement choices, then we expect greater weight on nonfinancial measures that are more highly associated with future financial performance.

Subjective versus Objective Measures

Prendergast and Topel's (1996) principal-agent model extends these studies to examine how favoritism and bias in subjective performance evaluations affect the relative weights placed on objective versus subjective performance measures in compensation and promotion decisions. They argue that subjectivity opens the door to favoritism, where supervisors act on personal preferences toward subordinates to favor some employees over others. Their model indicates that favoritism leads firms to place too little weight on supervisor appraisals and other subjective opinions of performance and too much weight on "hard" performance measures when combining multiple indicators in order to constrain favoritism and reduce the noise in the performance information. These results suggest that greater weight will be placed on more objective, quantitative measures than on more subjective, qualitative measures.

Psychology-Based Hypotheses

Driver versus Outcome Measures

Psychology-based studies diverge from the rational choice models in the economics literature to investigate how human information-processing limitations and decision strategies influence the use of performance measures. One of the most frequently observed results relates to the "outcome effect," in which evaluators systematically overweight outcome knowledge when assessing a manager's performance (e.g., Mitchell and Kalb 1981; Baron and Hershey 1988; Hawkins and Hastie 1990; Lipe 1993; Ghosh and Lusch 2000).

These studies find that evaluators tend to evaluate managers positively (negatively) when the outcome is positive (negative), regardless of whether the actions taken to achieve the results were appropriate. According to this literature, the overemphasis on outcomes arises because outcome knowledge influences the evidence recalled by the evaluator when attempting to assess the performance of the manager (Slovic and Fischhoff 1977). Consistent with these studies, we expect greater weight to be placed on outcome measures than on measures of the drivers of desired results, even when the driver measures are informative of the manager's actions.

In a similar vein, the balanced scorecard literature generally assumes that internal measures, such as innovation, process improvement, and employee satisfaction, are leading indicators or drivers of the outcome-oriented external (financial and customer) measures. If evaluators perceive external measures to be indicators of desired outcomes, then the outcome-effects literature suggests greater weight will be placed on external measures than on internal measures.

Financial versus Nonfinancial Measures

Relatively few psychology-based performance evaluation studies directly examine the relative weights placed on financial and nonfinancial measures for compensation purposes. Moreover, the results from related experimental studies provide mixed hypotheses regarding the relative weights placed on these measures. As discussed above, experimental evidence on outcome effects indicates that outcome or results-oriented measures will be weighted more heavily than input or driver measures. Since improved financial results are the ultimate goal of balanced scorecard systems (Kaplan and Norton 1996), outcome-effect studies also suggest that financial results will be weighted more heavily than nonfinancial results.

These studies also find that outcome effects are more pronounced when evaluators have used the outcome measures in earlier reward systems. Frederickson et al. (1999), for example, examine whether previous experience under a performance evaluation system systematically biases subsequent evaluations. They find that prior experience with an outcome-based performance evaluation system leads to larger outcome effects, with increased outcome feedback frequency increasing the effect. These results suggest that in a research setting such as ours, where financial results traditionally have been the primary performance objectives in bonus plans, greater weight will be placed on financial measures.

Lipe and Salterio (2000), in turn, examine whether evaluations using a balanced scorecard are affected by measures that are unique to a given organization, as well as measures that are common across organizations. Their experiment indicates that only common measures impact superior's evaluations, suggesting that measures that are used throughout an organization or are based on a common methodology will receive greater weight. To the extent that financial measures are more common and standardized across the organization's subunits, we expect greater weight on these measures.

More direct experiments on the use of financial and nonfinancial measures are inconclusive. Schiff and Hoffman (1996) find that executives tend to place greater weight on financial information when evaluating the performance of a business unit. However, they also find that participants place greater weight on nonfinancial information when evaluating a manager's performance. The authors provide no theoretical explanation for these differences.

Finally, Luft and Shields (2001) examine the use of financial and nonfinancial measures in a decision-making context. Their experiment indicates that participants place greater weight on current nonfinancial information when forecasting future financial performance than on current financial information. The authors attribute this result to nonfinancial measures being more cognitively valuable (i.e., more meaningful, transparent, and understandable) than financial measures. However, their experiment does not address the use of financial and nonfinancial measures for performance evaluation. The conflicting results in the preceding studies make hypotheses regarding the relative weights on financial and nonfinancial measures in bonus decisions unclear.

Objective/Quantitative Measures versus Subjective/Qualitative Measures

The organizational psychology literature has long held that greater weight should be placed on performance measures that are more reliable (e.g., Bellows 1954; Blum and Naylor 1968). According to this literature, subjective, qualitative performance assessments are often less accurate and reliable than more objective, quantitative measures because they are influenced by the rater's biases (e.g., Feldman 1981; Heneman 1986; Campbell 1990). Support for claims that subjective measures are less reliable than objective, quantitative measures is provided by studies examining the associations between subjective and quantitative measures covering the same activities. Meta-analyses by Heneman (1986) and Bommer et al. (1995) find correlations of only 0.27 and 0.39, respectively, between objective and subjective performance evaluations. Their analyses (as well as Heneman et al.'s [1987] review of related experimental studies) also indicate that differences between subjective and objective measures can be substantially reduced when the subjective ratings are based on the aggregation of multiple measures rather than on a single overall rating, and when performance is rated relative to a target or peer group rather than on an absolute scale. Assuming reliability is an important factor in the choice of performance measures, these studies suggest that greater weight will be placed on quantitative measures than on qualitative measures, and that greater weight will be placed on measures that are based on aggregations of multiple indicators and performance relative to targets than on other measures.

III. RESEARCH SETTING

A summary of our research hypotheses is presented in Table 1. We test these hypotheses using data from the North American retail banking operations of Global Financial Services (GFS), a leading international financial services provider. Prior to the 1990s, GFS had a strong financial orientation that focused almost exclusively on the achievement of financial results. Following a significant downturn in financial performance in the early 1990s, the bank initiated a review to evaluate its strategic direction. One conclusion of this review was that the bank's problems could be traced to a single-minded focus on bottom-line earnings and revenue growth to the exclusion of other important issues. This conclusion led to the development of a new retail banking business model in 1992 (see Figure 1). According to this model, the traditional goals of market share growth and financial performance were driven by more fundamental factors including product and service quality, cost effectiveness, risk control, employee relations, innovation, and customer satisfaction. The bank's executives agreed that organizational success would require greater emphasis on *all* of these dimensions.

Performance Incentive Plan

To support the new business model, GFS modified its bonus plans to emphasize the perceived drivers of profitability and growth. Until 1992, GFS awarded bonuses to branch managers based on a single branch profitability measure. Beginning in 1993, this plan was

TABLE 1
Summary of Hypotheses regarding the Relative Weights Placed on Different Types of Performance Measures

Economics-Based Hypotheses

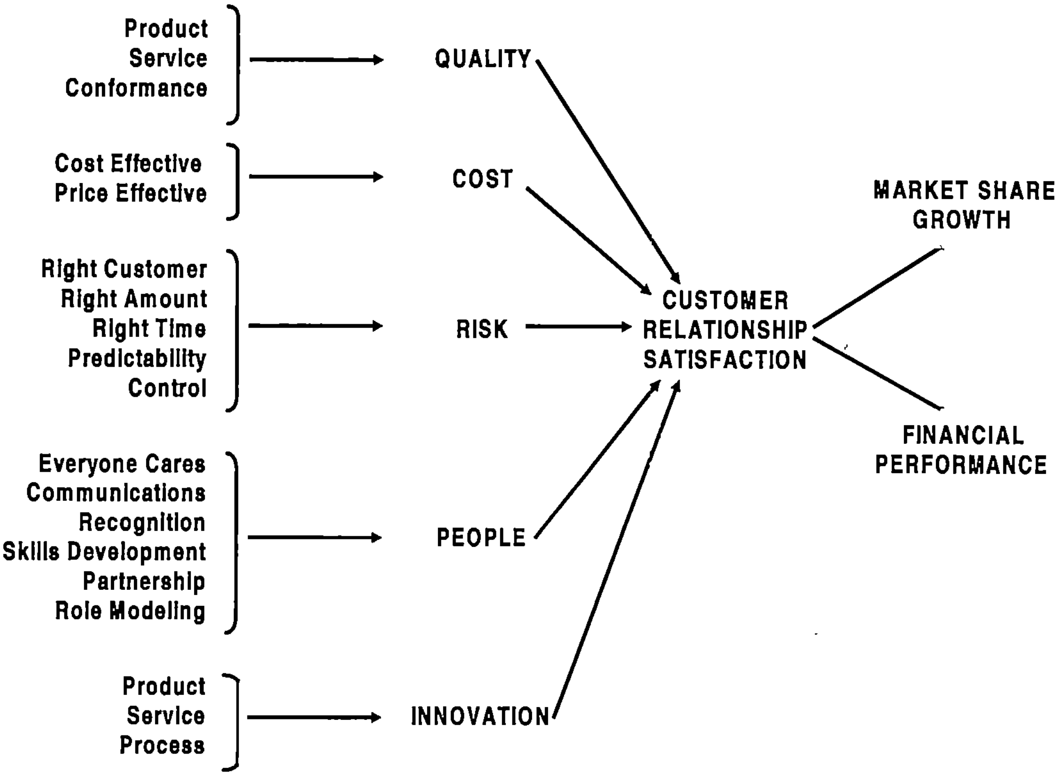
- Nonfinancial measures that are more predictive of financial results > Nonfinancial measures that are less predictive
- Objective measures > Subjective measures

Psychology-Based Hypotheses

- Outcome/results measures > Input/driver measures
- External measures > Internal measures
- Financial measures >/< Nonfinancial measures
- Objective/Quantitative measures > Subjective/Qualitative measures
- Measures based on multiple indicators > Single-item measures
- Measures with targets > Measures without targets

A > B: weight on measure A is greater than the weight on measure B.
A >/< B: weight on measure A has an ambiguous magnitude relative to the weight for measure B.

FIGURE 1
1992 Business Model for U.S./European Consumer Bank



replaced by the formula-based Performance Incentive Plan (PIP), which rewarded performance based on multiple, quantitative financial and nonfinancial measures. The PIP program's mechanics and evolution from 1993 to 1995 are summarized in Appendix A. Initially, branches were first required to receive satisfactory scores on any internal operational audits conducted during the quarter and to pass a customer satisfaction hurdle, as measured by a market research firm's survey of customer satisfaction with branch performance. Branches passing the customer satisfaction hurdle received quarterly bonuses for achieving improvement targets in any one of eight performance objectives related to growing the business (tier I and tier II household growth, consumer checking balance growth, business and professional checking balance growth, revenue growth, and relationship growth), resource management (expenses as a percent of revenue and footings as a percent of tier I and tier II households), and "overall performance" (quarterly margin growth).[3]

Minor changes occurred in 1994. In addition to passing the satisfaction hurdle and having satisfactory audit scores, branch managers were required to achieve targets in at least four of the eight performance objectives to receive a bonus. In 1995, the PIP objectives shifted further and included customer satisfaction (80 percent of customers rating overall satisfaction with GFS in the top two categories of the seven-point survey scale), growth (in tier I and tier II households, checking balances, liabilities and assets, and revenues), and resource management (growth in margins, and usage of automated tellers and other remote channels). To be eligible for bonuses, managers had to pass the customer satisfaction hurdle, have a satisfactory audit score, and meet their financial (revenue and margin) targets.

The growing complexity of the PIP bonus formula was reflected in the size of the document outlining each year's program: nine pages in 1993, 38 pages in 1994, and 78 pages in 1995. The primary cause of this complexity was upper management's frustration with a formula-based compensation system that allowed branch managers to game the system and earn bonuses without delivering financial results. To insure that branches were achieving financial targets, the 1995 PIP program added a financial hurdle that made it much more difficult for unprofitable branches to receive bonuses.

Balanced Scorecard Bonus Plan

Despite the increasing complexity of the PIP system, an internal evaluation concluded that results under the system had not put the organization on a trajectory to realize its strategic goals. In early 1995, the North American Banking Division (NABD) began replacing the Performance Incentive Plan with a "balanced scorecard" system focused on GFS's five corporate "imperatives" for success over time: achieving good financial results, delivering for customers, managing costs strategically, managing risk, and having the right people in the right jobs.[4] Scorecard implementation was accompanied by extensive training and widespread internal and external communication of the scorecard system's mechanics and objectives.

NABD's Western region replaced the PIP program with the balanced scorecard system in May 1995, with other NABD regions following in 1996. The research site provided us with extensive data on the Western region's balanced scorecard performance measures and performance evaluations, branch manager compensation, financial performance, and related internal employee surveys. Internal documents and employee survey data were also provided for other GFS regions. We supplemented this material with extensive interviews of GFS personnel at all organizational levels, and with observations of balanced scorecard working groups, implementation teams, and bonus award meetings over the life of the scorecard plan.

---

[3] A household is a group that makes banking decisions as a family or business unit. Tier I households are customers with total combined balances in excess of $100,000 (including investment balances) and tier II households are customers with balances in excess of $10,000. Footings are defined as consumer and business/professional liabilities plus consumer and business/professional assets (excluding mortgages).

[4] Implementation of the scorecard was not limited to NABD. Each business unit in GFS was required to develop a scorecard that focused on the five corporate imperatives. More importantly, the company's training material and internal communications indicated that the firm expected business units to achieve superior performance on all five dimensions, using the slogan "Five out of five, they all matter."

Balanced Scorecard Measures

The performance measures in the Western region's balanced scorecard fell into six categories: financial, strategy implementation, customer, control, people, and standards. The first three categories were each measured using multiple *quantitative* indicators. Financial performance was evaluated based on revenues, expenses, and margins. Depending upon the quarter, between seven and 18 measures were used to evaluate strategy implementation. These measures primarily related to growth in the number of customers in different segments, customer attrition, and the level of assets under management (AUM) for each customer or customer segment.

Two measures evaluated customer-related performance: overall satisfaction with GFS and the branch quality index.[5] Control was measured by the results of periodic internal audits of operations and legal/regulatory compliance. The people and standards evaluations represented *qualitative* assessments by the branch managers' area director. Superiors were instructed to consider performance management, teamwork, training and development (both for the branch manager and other branch employees), and employee satisfaction when assessing people-related performance.[6] Standards criteria were leadership, business ethics and integrity, customer interaction and focus, community involvement, and contribution to the overall business.[7] In Table 2, we summarize the characteristics of the measures in the scorecard categories and relate these measures to the research hypotheses in Table 1.[8]

Bonus Determination Process

The Western region was organized into five geographic areas, each consisting of five to 20 branches.[9] Branch managers within these areas reported to an area director, who reported to the president of the Western banking operation. Unlike the formula-based PIP program, the balanced scorecard system required area directors to *subjectively* weight the various performance measures when evaluating branch managers' performance and determining their bonuses. This process is summarized in Figure 2. Area directors first compared

---

[5] See Appendix B for the questions used to compute the two customer satisfaction measures.

[6] Performance management was defined as a manager's ability to "achieve goals by coaching, motivating, empowering, hiring, supporting, promoting, recognizing, and challenging staff." Although employee satisfaction was considered in evaluating the people category, employee satisfaction surveys were not conducted on a regular basis, making the quarterly assessment of this measure qualitative. Moreover, there was no statistically significant correlation between the employee satisfaction scores from a 1996 survey and the subjective "people" scores given by area directors in the first and second quarters of 1996, indicating that quantitative employee satisfaction measures received little weight in evaluating managerial performance on this dimension.

[7] Formal targets were not provided for the control, people, and standards categories, but an audit rating of "3" or lower was "below par" performance in the control category.

[8] Definitions for all the variables used in our statistical tests are provided in Appendix B.

[9] A total of 95 branches were open at some point during the period covered by our study. The number of branches open in an individual quarter ranged from 76 to 89. Missing balanced scorecard data reduce the sample sizes in some of our quarterly tests. This is particularly true for 1997, when balanced scorecard data were only available for 45, 37, 21, and 36 branches in the four quarters, respectively. The explanation from GFS regarding the missing data was that the scorecards were lost and could not be retrieved by the company. We have no way to assess the degree to which these lost observations bias our results.

TABLE 2

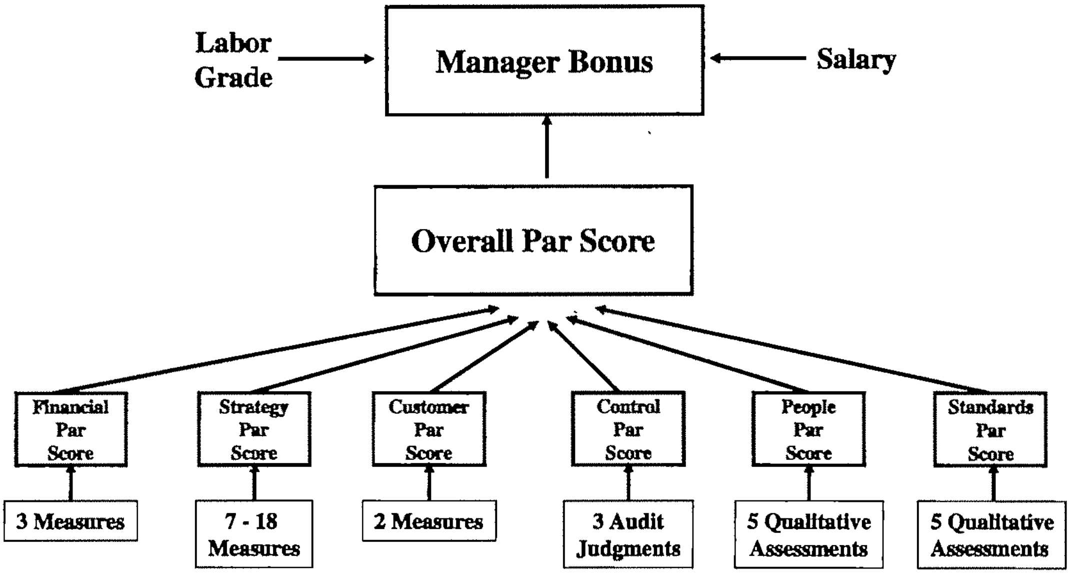
Characteristics of GFS's Balanced Scorecard Measures and Link to Measurement Characteristics Used in the Research Hypotheses

	Balanced Scorecard Category					
Measurement Characteristic	**Financial**	**Strategy**	**Customer**	**Control**	**People**	**Standards**
Financial	x	Some				
Objective/Quantitative	x	x	x			
Subjective				x	x	x
Outcome measure	x	x				
Externally oriented	x	x	x			
Targets set	x	Some[a]	x	Audit hurdle		
Multiple measures	x	x	x			
Single overall assessment				x^b	x^b	x^b

[a] Targets were set for all strategy implementation measures in the first four quarters under the balanced scorecard. In subsequent quarters, GFS set targets for only a subset of the strategy measures.

[b] Although area directors were instructed to consider multiple factors in assessing performance on these dimensions, only one overall assessment was given for each of these categories.

FIGURE 2
Balanced Scorecard Bonus Award Process



results to targets for each of the various financial, strategy implementation, and customer measures.[10] Branch managers then received a "par rating" for each of the measures within the financial, strategy, and customer categories, where "below par" reflected performance below expectations, "at par" represented expected performance, and "above par" reflected better-than-expected performance. Ratings for performance on the individual measures were then subjectively aggregated into ratings for the financial, strategy, and customer categories. The control, people, and standards categories each received a single *overall* par score (i.e., par scores were not given for the individual performance criteria within these three categories). The area director then subjectively combined the six scorecard categories into an overall performance rating of "below par," "at par," or "above par."

The quarterly bonuses recommended by the area directors were meant to reflect the branch manager's overall par score, labor grade, and current compensation. Unlike the PIP program, no formula was applied. Instead, bonuses were intended to achieve *total* market-based compensation levels (salary plus bonus) for a given labor grade and performance level. For example, assume that total compensation for branch managers in the highest of the three labor grades was targeted at *up to* \$75,000 annually if performance was at par, *up to* \$90,000 if performance was above par, and *up to* \$105,000 or more if performance was exceptional. If a manager with an above par overall evaluation in this labor grade earned a salary of \$80,000, then the *maximum* quarterly bonus was \$2,500 (\$10,000/4). However, if the manager's salary was \$90,000 or more, no bonus was awarded despite the above par performance. This differed from the PIP formula, which determined a bonus percentage that was awarded regardless of the branch manager's salary (e.g., a branch manager with an annual salary of $80,000 and eligible for 15 percent bonus would receive a $3,000 bonus for that quarter, while a branch manager with an annual salary of $90,000 and eligible for a 15 percent bonus would receive a quarterly bonus of $3,375).

---

[10] Targets were only provided for a subset of strategy implementation measures after the fourth quarter under the plan.

Area directors' bonus recommendations were taken to a meeting where the president of the Western region, the president's staff (the finance director, human resource director, compensation manager, and service quality director), and the five area directors discussed each recommendation. The discussion generally focused on the justification for the overall rating recommended for the branch manager, particularly when the overall evaluation of a manager was above par and the manager was eligible for a substantial bonus. Performance evaluations and bonuses could then be modified to promote consistency throughout the region.[11]

IV. RESULTS

Determinants of Scorecard Ratings

Our first set of tests addresses the weights placed on various types of performance measures when subjectivity is allowed in reward systems. We examine the *implicit* weights placed on the various performance measures each quarter by investigating the associations between the branches' performance on the scorecard measures, the par ratings given to branch managers in the six scorecard categories, and the size of their quarterly bonuses. These tests are similar to experimental studies using lens model methods to determine the relative weights placed on different types of information when evaluating performance, as well as archival compensation research that estimates the weights placed on accounting and stock market measures based on their statistical associations with compensation levels.

We first examine the regression weights placed on various *quantitative* measures.[12] Table 3 provides evidence on the associations between quarterly financial, customer, and strategy par ratings (where 1 = below par, 2 = at par, and 3 = above par) and performance versus targets for each of the quantitative performance measures in these categories.[13]

The financial par results in Panel A of Table 3 indicate that performance relative to revenue targets was factored into managers' evaluations during each of the quarters, with higher financial performance evaluations when revenues exceeded targets.[14] Expenses were negative in every quarter and statistically significant in ten of the 15 quarters ($p < 0.10$, two-tailed). In addition, the coefficients on revenues were substantially larger in absolute

---

[11] In contrast to the formula-based PIP system, which required no input by area directors, the scorecard process took area directors approximately six days per quarter to prepare the branch managers' scorecards, discuss the scorecards and bonus recommendations at the quarterly bonus meetings, and meet with the region's president. The extensive system of reviews under the scorecard plan is consistent with Prendergast and Topel's (1996) economic model, which indicates that extensive bureaucratic procedures are required to minimize the potential for favoritism and bias in subjective performance evaluations.

[12] For tests where the dependent variable is a par score (coded as 1, 2, or 3), we also conducted the statistical analysis using an ordered regression. These results were virtually identical to those reported in the tables.

[13] As noted earlier, GFS did not set targets for *all* of the strategy implementation measures after the first four quarters of the balanced scorecard. Consequently, our strategy par score analyses use absolute performance on measures without a goal for the later quarters and relative performance when goals are available. For measures with goals, we use the ratio of the actual outcome to the goal as the predictor variable.

[14] We delete observations from all regression tests when studentized residuals are greater in absolute value than 3 to mitigate the impact of outliers on our results.

TABLE 3
Regression Models Examining the Relation between Financial and Customer Par Scores and Their Underlying Quantitative Performance Measures

Par Score = f(Quantitative Performance Measures)

Panel A: Financial Par Scores

	Second Quarter 1995	Third Quarter 1995	Fourth Quarter 1995	First Quarter 1996	Second Quarter 1996	Third Quarter 1996	Fourth Quarter 1996	First Quarter 1997	Second Quarter 1997	Third Quarter 1997	Fourth Quarter 1997	First Quarter 1998	Second Quarter 1998	Third Quarter 1998	Fourth Quarter 1998
Revenue	8.01***	8.51***	4.73***	6.76***	5.30***	0.44*	7.00***	8.12***	6.44***	4.74***	4.03***	1.52***	1.26***	4.69***	3.01***
Expense	−1.26***	−2.14***	−4.32***	−2.56***	−1.61**	−0.77	−3.23***	−2.21**	−1.38	−3.17**	−0.80**	−0.36	−0.09	−0.91	−0.61**
Adj. R^2	0.64	0.57	0.49	0.56	0.49	0.05	0.62	0.57	0.37	0.49	0.30	0.18	0.14	0.41	0.31
n	73	74	83	89	89	81	80	45	37	21	36	74	81	66	58

Panel B: Customer Par Scores

	Second Quarter 1995	Third Quarter 1995	Fourth Quarter 1995	First Quarter 1996	Second Quarter 1996	Third Quarter 1996	Fourth Quarter 1996	First Quarter 1997	Second Quarter 1997	Third Quarter 1997	Fourth Quarter 1997	First Quarter 1998	Second Quarter 1998	Third Quarter 1998	Fourth Quarter 1998
GFS CSI	6.73***	7.83***	7.40***	7.64***	10.73***	9.02***	2.55***	7.60***	6.75***	7.52***	7.70***	7.78***	6.63***	6.35***	7.63***
Branch Quality Index	1.49***	0.41	−0.17	0.30	−0.84**	−0.41	1.15*	−0.17	−0.04	0.88*	0.19	−0.46	0.96**	0.17	−0.14
Adj. R^2	0.76	0.70	0.78	0.66	0.80	0.74	0.21	0.83	0.78	0.76	0.77	0.69	0.75	0.70	0.67
n	73	74	79	62	84	80	79	43	86	82	72	78	59	66	58

***, **, * Statistically significant at the 1%, 5%, and 10% levels, respectively (two-tailed).
See Appendix B for variable definitions.

value than those on expenses, suggesting that revenues received greater weight in performance evaluations. The mean (median) adjusted R^2 was 0.41 (0.49), implying that the quantitative performance measures for this category accounted for a little less than half of the managers' performance evaluation on this dimension.[15]

Nearly all the emphasis in customer-related evaluations was on overall satisfaction with GFS (Table 3, Panel B). The overall customer satisfaction measure was a significant determinant of customer par ratings in every quarter. The branch quality index, on the other hand, had a significant positive impact on customer par ratings in only four of the 15 quarters, and was significantly *negative* in the second quarter of 1996. Our interviews indicated that the primary reason for the heavy weight on overall satisfaction rather than the branch quality index was the fact that *all* senior managers and business units in GFS were evaluated using the common overall customer satisfaction measure. As a result, the Western region's president, who was also held responsible for this measure, drove this measure down to lower organizational levels in his division in order to provide an assessment of the region's progress toward the corporate goal of 80 percent overall satisfaction that was established for all GFS businesses. This result is consistent with Lipe and Salterio's (2000) experimental finding that evaluators place greater weight on common balanced scorecard measures than on unique measures, as well as contagion theories that suggest that lower-level managers evaluate subordinates using the same criteria used by upper-level managers to evaluate their performance (e.g., Hopwood 1974). In general, the customer par score models had substantially greater explanatory power than the financial models (mean adjusted $R^2 = 0.71$, median = 0.75)

Determinants of the strategy par scores are examined in Table 4. Given the large number of strategy performance measures (7 to 18 depending upon the quarter) and the small sample sizes in some of the quarters, we use stepwise regression models in these tests. The explanatory power of the strategy implementation measures varied widely from quarter to quarter, with adjusted R^2s ranging from 0.10 in the first quarter of 1996 to 0.72 in the third quarter of 1997 (mean = 0.32, median = 0.27). Moreover, no more than five of the measures were significant predictors of strategy par scores in any quarter ($p < 0.15$, two-tailed), and in many quarters only one or two of the strategy measures were significant (mean = 2.73 categories, median = 2).

The results in Tables 3 and 4 indicate that the subjective evaluations given to branch managers for the financial, customer, and strategic implementation categories were based only partially on the quantitative performance measures they were supposed to reflect. Adjusted R^2s ranged from 0.05 to 0.81, indicating that roughly 19 to 95 percent of a branch manager's par rating for a particular category was based on factors other than performance relative to the category's goals. Obviously, the explanatory power of these models is subject to the validity of the model specification linking the independent and dependent variables. Despite this caveat, the large unexplained variance in many of the models is consistent with analytical research indicating that the presence of subjectivity and discretion in reward systems allows superiors to base evaluations on factors other than the measures included in the compensation plan (Baker et al. 1994; Baiman and Rajan 1995; Prendergast and Topel 1996), even though this discretion was not intended by the scorecard plan's designers.

---

[15] We did not include margins (defined as revenues – expenses) in the model because of multicollinearity problems. When financial par scores were regressed on margins alone, the coefficients were positive and significant in each quarter. However, the adjusted R^2s for the models were substantially lower (mean = 0.22, median = 0.21). Low variance inflation factor (VIF) scores indicate no serious problems with multicollinearity in any of the other models reported in the paper.

TABLE 4
Stepwise Regressions Examining the Relation between Strategy Implementation Par Scores and Their Underlying Quantitative Performance Measures

Par Score = f(Quantitative Performance Measures)

Second Quarter 1995	Total AUM ($)^R	4.62***
Adj. $R^2 = 0.41$	Retail Households (#)^R	6.02***
$n = 48$	Household Attrition (#)^R	−1.74***
Third Quarter 1995	B&P Households (#)^R	3.11***
Adj. $R^2 = 0.12$	Total AUM ($)^R	2.18*
$n = 57$		
Fourth Quarter 1995	Retail Households (#)^R	6.92***
Adj. $R^2 = 0.27$	Total AUM ($)^R	0.93***
$n = 73$		
First Quarter 1996	B&P Households (#)^R	0.94***
Adj. $R^2 = 0.10$		
$n = 75$		
Second Quarter 1996	Retail Households (#)^R	0.42***
Adj. $R^2 = 0.29$	B&P CNR/Household ($)	0.00*
$n = 58$		
Third Quarter 1996	Total Cross Sales ($)	0.01***
Adj. $R^2 = 0.34$	B&P Households (#)^R	0.14*
$n = 41$		
Fourth Quarter 1996	New Premier Households (#)	0.04***
Adj. $R^2 = 0.26$	Retail Households (#)^R	0.27**
$n = 46$		
First Quarter 1997	Premier CNR/Household ($)^R	0.56**
Adj. $R^2 = 0.41$	Lost Retail Households (#)	−0.01***
$n = 43$	New Retail Households (#)	0.01**
Second Quarter 1997	B&P CNR/Household ($)^R	1.69***
Adj. $R^2 = 0.63$	New Premier Households (#)	0.13***
$n = 36$	Lost B&P Households (#)	−0.04***
	New B&P Households (#)	0.02**
	Retail Households (#)^R	0.94*
Third Quarter 1997	Automated Transactions (% of total)	0.03***
Adj. $R^2 = 0.72$	Premier CNR/Household (#)^R	0.81***
$n = 21$	Lost B&P Households (#)	−0.04***
	New Retail Households (#)	0.01**
	B&P CNR/Household ($)^R	0.79*
Fourth Quarter 1997	New Premier Households (#)	0.05***
Adj. $R^2 = 0.24$		
$n = 25$		
First Quarter 1998	Retail Households (#)^R	1.64***
Adj. $R^2 = 0.23$	New Retail Households (#)	0.03***
$n = 70$	Lost Retail Households (#)	−0.00*
Second Quarter 1998	New B&P Households (#)	0.06***
Adj. $R^2 = 0.32$	Retail Households (#)^R	1.21***
$n = 52$	Lost Retail Households (#)	−0.00#
	New Retail Households (#)	0.01*
	Premier Households (#)^R	0.52*
Third Quarter 1998	Premier CNR/Household ($)^R	0.71***
Adj. $R^2 = 0.23$	Premier Households (#)^R	0.49**
$n = 59$		
Fourth Quarter 1998	Premier CNR/Household ($)^R	0.35***
Adj. $R^2 = 0.22$	B&P CNR/Household ($)^R	0.91***
$n = 53$	Retail Households (#)^R	1.18*

***, **, *, # Statistically significant at the 1%, 5%, 10%, and 15% levels, respectively (two-tailed).

^{R} Performance measure is relative to goal.

See Appendix B for variable definitions.

Association between Ratings in Individual Scorecard Categories and Overall Ratings

Our second set of tests investigates the association between the qualitative ratings given in each of the six scorecard categories and the branch manager's overall performance rating. Since the par scores were categorical variables (below par, par, or above par), we estimate general linear models with indicator variables for the three par score categories. We evaluate the proportion of variance in the dependent variable attributable to each par score using partial eta squareds (η^2).[16]

The results in Table 5 indicate that the financial and customer par scores were the most consistent determinants of overall performance ratings. Partial eta squareds for the financial par scores were significant in all but one of the quarters ($p < 0.10$, two-tailed F-test), while customer par scores were significant in 12 of the quarters. In contrast, the qualitative standards, people, and control par scores were only significant determinants of overall par scores in five, four, and two quarters, respectively. Surprisingly, the strategy par score, which was intended to be based on quantitative performance indicators, was only significant in six of the 15 quarters, a result similar to the qualitative standards assessment. This evidence is inconsistent with our hypothesis that quantitative performance measures or assessments based on multiple indicators receive greater weight in performance evaluations, but is consistent with studies finding lower weight on measures when performance targets (which were not set for all of the strategy measures after the first four quarters under the plan) are not provided (e.g., Heneman 1986; Bommer et al. 1995).

To examine whether the magnitudes of the partial eta squareds were significantly different among the performance categories, we compute nonparametric Wilcoxon signed rank tests between each pair of categories (not reported in the tables). Median eta squareds for the financial par scores were significantly larger than those for any of the other performance categories ($p < 0.01$, two-tailed). Median eta squareds for the customer par scores were also significantly larger than medians for strategy implementation and the three qualitative categories. Strategy implementation eta squareds, on the other hand, were only marginally larger than those for the control and standards categories ($p < 0.15$, two-tailed), and were not significantly different than the people eta squareds.

These analyses indicate that financial par scores were used more frequently and received greater weight than nonfinancial par scores. As hypothesized, both the financial and customer par scores, which were based on externally oriented, quantitative results measures, received greater emphasis than the more qualitative, internally oriented customer, people, and control par scores. However, the limited emphasis placed on strategy par scores is inconsistent with our hypotheses that quantitative measures and subjective evaluations based on multiple criteria receive greater weighting. Although strategy par scores were meant to reflect their underlying quantitative measures (which the research site's business model assumed to be key externally oriented outcome measures), we find that they were significant determinants of overall par scores in only 40 percent of the quarters, and generally received weights similar to the single-item qualitative performance evaluations.

Finally, even though nonfinancial performance was incorporated into the area director's overall assessment of a branch manager, the significantly larger eta squareds for the financial par scores indicate that the balanced scorecard system was primarily driven by financial

---

[16] The partial eta squared is the proportion of the effect plus error variance that is attributable to an effect, and is calculated as: $\eta_p^2 = SS_{effect}/(SS_{effect} + SS_{error})$. The sums of the partial eta squared values are not additive, and do not sum to the amount of dependent variable variance accounted for by the independent variables. Furthermore, it is possible for the sums of the partial eta squareds to be greater than 1.

TABLE 5
Partial Eta Squareds from GLM Models Examining the Relation between Overall Par Scores and Individual Par Scores

Overall Par Score = f(Individual Par Scores)

	Second Quarter 1995	Third Quarter 1995	Fourth Quarter 1995	First Quarter 1996	Second Quarter 1996	Third Quarter 1996	Fourth Quarter 1996	First Quarter 1997	Second Quarter 1997	Third Quarter 1997	Fourth Quarter 1997	First Quarter 1998	Second Quarter 1998	Third Quarter 1998	Fourth Quarter 1998
Financial	0.42***	0.28***	0.56***	0.63***	0.53***	0.37***	0.21***	0.44***	0.66***	0.35	0.33*	0.43***	0.37***	0.30***	0.41***
Strategy	0.02	0.14***	0.13***	0.08	0.12**	0.01	0.02	0.04	0.20	0.01	0.05	0.28***	0.29***	0.26**	0.12
Customer	0.11**	0.10**	0.06	0.24***	0.50***	0.24***	0.16***	0.01	0.34**	0.27	0.36*	0.30***	0.33***	0.38***	0.24*
Control	0.02	0.06**	0.02	0.32***	0.09*	0.01	0.01	0.18*	0.00	0.05	0.09	0.12*	0.14*	0.09	0.00
People	0.00	0.14***	0.04	0.16**	0.06	0.08	0.27***	0.07	0.01	0.13	0.03	0.30***	0.01	0.24**	0.09
Standards	0.08*	0.06	0.08**	0.02	0.01	0.08*	0.01	0.27**	0.02	0.06	0.00	0.05	0.09*	0.02	0.02
Model Adj. R^2	0.50	0.55	0.75	0.87	0.79	0.60	0.53	0.74	0.72	0.38	0.65	0.67	0.75	0.73	0.71
n	73	72	80	61	70	53	69	42	26	20	25	56	41	42	31

***, **, * Statistically significant at the 1%, 5%, and 10% levels, respectively (two-tailed).
See Appendix B for variable definitions.

considerations. This evidence suggests that the balanced scorecard plan only partially addressed criticisms that GFS's earlier formula-based compensation program provided incentives for branch managers to focus their efforts on a single or limited set of performance dimensions.

Associations between Scorecard Ratings and Quarterly Bonuses

We examine the extent to which the performance ratings in the six scorecard categories were weighted in bonus computations in Table 6. Because the maximum bonus award under the scorecard was intended to be a function of the branch manager's labor grade and current salary, we include these factors as control variables in the analyses (where 1 = the lowest labor grade and 3 = the highest).

The partial eta squareds in Panel A of Table 6 investigate the relation between bonuses and branch managers' overall performance ratings. Results for labor grade and salary are not presented to simplify presentation.[17] The overall par score is a significant predictor of bonuses in every quarter ($p < 0.10$. two-tailed), with the models' explanatory power ranging from 0.19 to 0.70 (mean = 0.48, median = 0.45).

The influence of the individual scorecard categories on bonus awards is examined in Panel B of Table 6. Financial performance had the most frequent direct association with bonus awards. The financial par score was significant in nine of the 15 quarters, with a mean (median) partial eta squared of 0.27 (0.29). The other par scores were only significant in four or fewer quarters. This evidence suggests that any effects of nonfinancial measures on bonus awards were primarily through their influence on overall par scores, rather than through their direct influence on bonuses.

Taken together, the results in Tables 5 and 6 suggest that bonus computations under the balanced scorecard system were focused much more on financial outcomes than were bonus computations under the formula-based PIP system. Under the PIP system, the performance measures in the compensation plan could not be ignored as long as the manager met the appropriate hurdles and targets. In contrast, the results in Tables 3 and 4 indicate that a number of the performance measures in the scorecard had little or no influence on performance evaluations and bonus awards. Moreover, unlike the strict formula used to compute bonuses under the PIP system, the balanced scorecard performance measures explained only about 50 percent of the observed bonuses. The large unexplained variance in bonus amounts is consistent with the greater discretion in the balanced scorecard system than in the PIP system, which did not allow deviations from the bonus formula.

Associations between Nonfinancial Measures and Financial Performance

One potential explanation for the limited weights placed on some of the nonfinancial measures is that these measures contained little incremental information on managerial actions. Many economics-based agency models indicate that performance measures should only be included in compensation plans when they provide incremental information on the actions desired by the principal. A key assumption of GFS's scorecard system was that the nonfinancial scorecard measures were *leading* indicators of future financial results. If some of the nonfinancial balanced scorecard measures actually had no ability to predict future financial performance, their omission in performance evaluations could be due to area directors ignoring or abandoning uninformative performance indicators.

---

[17] As expected, bonuses were smaller at lower labor grades and at higher salary levels (after controlling for labor grade) due to the maximum targeted total compensation levels established for each labor grade.

TABLE 6
Partial Eta Squareds from GLM Models Examining the Relations between Bonus Awards and Par Scores

Panel A: Bonus Awards = *f*(Overall Par Score, Salary, Labor Grade)

	Second Quarter 1995	Third Quarter 1995	Fourth Quarter 1995	First Quarter 1996	Second Quarter 1996	Third Quarter 1996	Fourth Quarter 1996	First Quarter 1997	Second Quarter 1997	Third Quarter 1997	Fourth Quarter 1997	First Quarter 1998	Second Quarter 1998	Third Quarter 1998	Fourth Quarter 1998
Overall Par Score	0.24***	0.69***	0.33***	0.57***	0.44***	0.24***	0.34***	0.34***	0.62***	0.12	0.42***	0.29***	0.62***	0.15**	0.25***
Model Adj. R^2	0.38	0.70	0.61	0.60	0.55	0.38	0.33	0.31	0.67	0.45	0.45	0.32	0.62	0.19	0.57
n	56	64	63	61	75	69	73	41	31	18	35	70	57	62	56

Panel B: Bonus Awards = *f*(Individual Par Scores, Salary, Labor Grade)

	Second Quarter 1995	Third Quarter 1995	Fourth Quarter 1995	First Quarter 1996	Second Quarter 1996	Third Quarter 1996	Fourth Quarter 1996	First Quarter 1997	Second Quarter 1997	Third Quarter 1997	Fourth Quarter 1997	First Quarter 1998	Second Quarter 1998	Third Quarter 1998	Fourth Quarter 1998
Financial	0.30***	0.20**	0.47***	0.39***	0.32***	0.45***	0.23***	0.17	0.51**	0.29	0.21	0.09	0.10	0.02	0.36**
Strategy	0.10	0.23*	0.08	0.04	0.12*	0.06	0.05	0.09	0.06	0.41	0.24	0.06	0.21**	0.12	0.03
Customer	0.08	0.09	0.06	0.03	0.02	0.00	0.13**	0.01	0.15	0.12	0.46**	0.07	0.17*	0.02	0.14
Control	0.00	0.01	0.01	0.02	0.03	0.08	0.09	0.00	0.15	0.44	0.15	0.08	0.07	0.08	0.09
People	0.01	0.06	0.00	0.02	0.01	0.12*	0.03	0.08	0.02	0.02	0.08	0.01	0.08	0.01	0.09
Standards	0.02	0.02	0.10**	0.00	0.03	0.00	0.04	0.15	0.00	0.05	0.09	0.00	0.00	0.05	0.07
Model Adj. R^2	0.41	0.37	0.62	0.56	0.63	0.59	0.40	0.39	0.52	0.54	0.58	0.21	0.54	0.06	0.62
n	49	55	61	49	64	48	63	40	23	17	25	54	41	42	31

***, **, * Statistically significant at the 1%, 5%, and 10% levels, respectively (two-tailed).
Salary and labor grade results are not presented to simply presentation.
See Appendix B for variable definitions.

We provide some evidence on the informativeness of the nonfinancial measures by examining whether changes in these measures were leading indicators of the bank's two strategic objectives (customer growth and financial performance). Consistent with the business model in Figure 1, we regress percentage changes in financial performance or customer growth on percentage changes in the quantitative customer measures and the qualitative people, control, and standards par scores. We also include the customer growth measures as predictor variables in the financial performance models to examine whether changes in customer growth were associated with current or future changes in financial results. Indicator variables for time period are included as control variables to account for potential time-specific effects on performance.[18]

We examine these relations using three time lags. The first two tests examine short-term effects using contemporaneous changes and one quarter lags (e.g., percentage changes in the dependent variables between quarters 2 and 3 are regressed on percentage changes in the predictor variables between quarter 1 and 2). Although these lags are relatively short, the frequent repurchase cycle and relatively low customer switching costs in retail banking can lead to reasonably short lags between managerial actions and observed economic performance. Longer-term effects are examined using rolling four quarter lags. For example, percentage changes in financial performance and customer growth are computed using the current quarter in a given year and the same quarter one year earlier (e.g., quarter 1 results for 1998 divided by quarter 1 results for 1997). These changes are regressed on rolling percentage changes in the predictor variables during the previous four quarters (e.g., quarter 1 results for 1997 are divided by quarter 1 results for 1996). By computing percentage changes relative to the same quarter in the prior year, we control for seasonality.

The results are presented in Table 7. We find little evidence consistent with the hypothesis that the weights placed on nonfinancial measures are a function of their ability to predict future changes in performance (our proxy for informativeness). Although overall satisfaction with GFS was the primary determinant of customer par scores, this measure had no contemporaneous relation with customer growth or financial performance, and was *inversely* related to some of the future results. In particular, changes in GFS satisfaction were negatively related to changes in Retail and Premier customers and positively related to expense changes in the following quarter, and were negatively related to changes in Retail customers and margins in the subsequent four quarters. The associated reductions in Retail customers also appear to have had a negative *indirect* effect on margins due to the positive contemporaneous and four quarter lagged relations between Retail customers and margins and the positive relation with expense changes in the following quarter.

In contrast, the branch quality index, which received relatively little weight in the determination of customer par scores, had a significant positive association with four quarter lagged changes in margins but not with short-term margin changes, suggesting that improvements in branch satisfaction only yielded improvements in financial performance with some delay. The branch quality index was also positively associated with contemporaneous changes in retail customers, and negatively associated with expense changes in the subsequent quarter. Percentage changes in the number of retail customers, in turn, had positive relations with current revenues and margins and with four quarter lagged margins. Taken together, these results are consistent with branch quality having positive direct effects on

---

[18] We also estimated these models with indicator variables for each of the five distinct regions in an attempt to control for potential regional differences. These results were virtually identical to those reported in Table 7.

TABLE 7

Regression Models Examining the Relation between Percentage Changes in Customer Growth or Financial Performance and Percentage Changes in Nonfinancial Performance Measures

%ΔCustomer Growth or *%ΔFinancial Performance* = *f*(*%ΔNonfinancial Performance Measures, Time Period Indicators*)

Panel A: Contemporaneous Relations

	%ΔRetail Customers	*%ΔB&P Customers*	*%ΔPremier Customers*	*%ΔRevenue*	*%ΔExpenses*	*%ΔMargin*
Intercept	0.933***	1.123***	1.616***	0.833***	1.092***	0.443**
%ΔGFS CSI	−0.015	−0.026	0.099	0.025	−0.076	0.121
%ΔBranch Quality Index	0.086**	0.004	−0.026	0.006	−0.011	−0.058
%ΔControl Par Score	0.019	0.010	0.021	−0.019*	0.020	−0.022
%ΔPeople Par Score	−0.012	−0.038	−0.100*	0.043**	−0.045**	0.148***
%ΔStandards Par Score	−0.009	0.035	0.045	−0.011	0.000	−0.074
%ΔRetail Customers				0.113**	0.003	0.413***
%ΔB&P Customers				0.004	0.008	0.015
%ΔPremier Customers				0.006	0.012	0.009
Adj. R^2	0.058	0.006	0.235	0.096	0.309	0.182
F-statistic	2.583***	1.145	8.154***	3.129***	9.998***	5.419***
n	461	460	420	421	424	417

Panel B: One Quarter Lag

	%ΔRetail Customers	*%ΔB&P Customers*	*%ΔPremier Customers*	*%ΔRevenue*	*%ΔExpenses*	*%ΔMargin*
Intercept	1.189***	1.219***	1.393***	0.995***	1.102***	1.007***
%ΔGFS CSI	−0.093*	−0.068	−0.275*	−0.018	0.118*	−0.135
%ΔBranch Quality Index	−0.052	−0.053	0.039	0.031	−0.086*	0.111
%ΔControl Par Score	−0.015	0.008	−0.013	0.017	−0.017	−0.006
%ΔPeople Par Score	0.006	0.013	−0.011	−0.001	0.054***	−0.079*
%ΔStandards Par Score	0.015	−0.025	0.059	0.016	0.011	0.087

(*continued on next page*)

TABLE 7 (continued)

Panel B: One Quarter Lag (continued)

	%ΔRetail Customers	***%ΔB&P Customers***	***%ΔPremier Customers***	***%ΔRevenue***	***%ΔExpenses***	***%ΔMargin***
%ΔRetail Customers				0.003	0.011*	−0.006
%ΔB&P Customers				−0.003	−0.004	−0.006
%ΔPremier Customers				−0.007	0.027***	−0.025#
Adj. R^2	0.054	0.019	0.241	0.071	0.454	0.110
F-statistic	2.387***	1.462*	8.134***	2.420***	16.516***	2.997***
n	412	412	382	372	374	372

Panel C: Four Quarter Lag

	%ΔRetail Customers	***%ΔB&P Customers***	***%ΔPremier Customers***	***%ΔRevenue***	***%ΔExpenses***	***%ΔMargin***
Intercept	1.148***	1.283**	0.458	1.163*	1.317***	0.879***
%ΔGFS CSI	−0.136*	−0.169	0.692	−0.151	0.036	−0.438#
%ΔBranch Quality Index	−0.007	0.092	0.128	0.071	−0.107	0.411**
%ΔControl Par Score	−0.009	−0.046	0.096	−0.021	0.024	−0.076
%ΔPeople Par Score	−0.009	−0.002	0.395	−0.031	0.061**	−0.132**
%ΔStandards Par Score	0.002	0.024	0.191	−0.057	−0.019	−0.091
%ΔRetail Customers				0.107	−0.029	0.452**
%ΔB&P Customers				0.115**	−0.024	0.117
%ΔPremier Customers				0.001	0.004	−0.002
Adj. R^2	0.061	0.045	0.054	0.088	0.345	0.089
F-statistic	2.217***	1.892**	2.079***	2.331***	8.211***	2.333**
n	207	211	209	194	193	192

***, **, *, # Statistically significant at the 1%, 5%, 10%, and 15% levels, respectively (two-tailed).
Indicator variables for time period are included in the models but not reported to simplify presentation.
See Appendix B for variable definitions.

future margins, as well as positive indirect effects on current and future financial performance through growth in retail customers.

The other two customer-growth measures show little positive association with financial performance. B&P customer growth was associated with *lower* margins in the following month. Similarly, the percentage change in the number of Premier households, which was a significant determinant of strategy par scores in many quarters, was also *negatively* associated with margins in the following month due to the higher expenses to acquire and maintain these customers. Neither Premier nor B&P customer growth was associated with changes in financial performance over the next four quarters, despite the fact that measures related to these customer groups were significant determinants of strategy par scores in 12 of the 15 quarters.

Finally, changes in the qualitative control and standards par scores exhibit few significant associations with current or future performance, even though these measures had some influence on overall par scores. People par score changes had the expected positive associations with contemporaneous financial performance (higher revenue, lower expenses, and higher margins), but were *negatively* associated with subsequent (one-quarter and four-quarters ahead) margins due to higher expenses.

Although the results in Table 7 are subject to a variety of econometric concerns such as correct model specification, they are generally inconsistent with informativeness theories on the choice of performance measures. In the case of the branch quality index, the limited use of this measure appears to reflect the omission of value-relevant information from performance evaluations rather than the exclusion of a measure that provided no information on managerial performance. In other cases, measures that did not exhibit the statistical associations with financial performance predicted by GFS's business model received significant weight in performance evaluations. This is particularly true of the overall GFS customer satisfaction measure, which was *negatively* associated with the bank's strategic objectives. Together with the earlier results, this evidence suggests that psychology-based explanations may be equally or more relevant than economics-based explanations in explaining measurement practices at GFS.

V. QUALITATIVE ANALYSES

A summary of the preceding hypothesis tests is presented in Table 8. These statistical tests indicate that the subjectivity in the balanced scorecard plan allowed area directors to incorporate factors other than the scorecard measures in performance evaluations, to change evaluation criteria from quarter to quarter, to ignore measures that were predictive of future financial performance, and to weight measures that were not predictive of desired results. However, these analyses shed little light on whether the resulting scorecard plan achieved its objectives.

Some evidence on the outcomes of the balanced scorecard program can be obtained from two internal branch manager surveys conducted by GFS's human resource department.[19] Table 9 examines the balanced scorecard's influence on Western branch managers' perceived understanding of strategic goals and their attitudes toward performance evaluation criteria using data from internal employee surveys conducted in October 1994 (under the

---

[19] Sample sizes were 77 for the 1994 survey and 83 for the 1996 survey. Response rates were 97 percent and 93 percent, respectively.

TABLE 8
Summary of Hypothesis Test Results

Hypothesis regarding Relative Weights	Results
Economics-Based Hypotheses	
Nonfinancial measures that are more predictive of financial results > Nonfinancial measures that are less predictive	Not consistent (Table 7)
Objective measures > Subjective measures	Consistent for financial and customer measures; not consistent for strategy measures (Tables 5 and 6)
Psychology-Based Hypotheses	
Outcome/Results measures > Input/Driver measures	Consistent for financial measures; not consistent for strategy measures (Tables 5 and 6)
External measures > Internal measures	Consistent for financial and customer measures; not consistent for strategy measures (Tables 5 and 6)
Financial measures >/< Nonfinancial measures	Greater weight on financial measures (Tables 5 and 6)
Objective/Quantitative measures > Subjective/Qualitative measures	Consistent for financial and customer measures; not consistent for strategy measures (Tables 5 and 6)
Measures based on multiple indicators > Single-item measures	Consistent for financial and customer measures; not consistent for strategy measures (Tables 5 and 6)
Measures with targets > Measures without targets	Consistent (Tables 5 and 6)

PIP program) and February 1996 (under the balanced scorecard system). The mean responses in 1994 and 1996 reveal few statistical differences in perceptions under the formula-based PIP program and the more subjective, but broader, balanced scorecard. The scorecard's implementation brought little change in branch managers' stated understanding of strategic goals or their connection to the managers' actions. Under both systems, branch managers, on average, claimed that they understood GFS's business goals, the goals of their work group, the connection between their jobs and the business objectives, and the basis on which performance was judged. In contrast, the managers generally agreed that the GFS strategy had become clearer to them between 1995 and 1996, suggesting that the scorecard may have helped to communicate the company's strategic goals. However, branch managers felt less comfortable with the adequacy of the information provided to them about progress toward the multiple business goals. The perceived importance of customers and employee development in performance evaluation and compensation decisions, as well as agreement with the statement "compensation decisions are consistent with performance," were not significantly different under the two systems.

The responses in Table 9 provide little evidence that the change from the PIP system to the balanced scorecard had an impact (either positive or negative) on managerial perceptions of business strategies, goals and priorities, performance evaluation, and compensation bases, or the adequacy of measures for decision making. However, the survey was conducted only three full quarters after the start of the scorecard system and was not focused on assessing the scorecard's impact. A subsequent survey covering a broader sample of

TABLE 9
Mean Employee Survey Responses by Western Region Branch Managers under the Performance Incentive Plan and Balanced Scorecard

	Performance Incentive Plan 1994 (n = 77)	**Balanced Scorecard 1996 (n = 83)**
A. Understanding of Strategy and Business Objectives		
I understand the business goals of GFS[a]	1.83	1.75
During the past year, the GFS strategy has become clearer to me[a]	NA	2.02
Senior management has communicated a clear plan for meeting our business goals[a]	2.13	2.32
I see the connection between the business objectives and my job[a]	1.93	1.71
B. Goal Setting		
I understand the goals of my work group[a]	NA	1.77
I know the basis on which my performance will be judged[a]	1.84	1.87
I get adequate information about progress against business goals[a]	2.00	2.87*
C. Performance Evaluation and Compensation		
Service to the customer is an important part of the way my performance is measured[a]	1.75	1.79
I am recognized for the service I provide to customers[a]	2.18	2.48
Managers are rewarded for developing their employees[a]	2.72	2.90
Decisions about my compensation have been consistent with my performance[a]	2.82	2.85

* Significantly different from the mean 1994 survey response at the 10% level or better (two-tailed).

[a] 1 = strongly agree, 2 = agree, 3 = neither agree nor disagree, 4 = disagree, 5 = strongly disagree.

North American managers was conducted in 1997 to specifically address the implementation of the scorecard.[20] Respondents were asked to use a three-point scale (agree, neutral, or disagree) to answer questions on their understanding of the scorecard process, implementation and goal setting, performance evaluation and bonus awards, and overall assessment of the scorecard process.

The results from the second survey are presented in Table 10. Although the majority of managers stated that they understood the scorecard process and their scorecard goals, only 32 percent of respondents were satisfied with the scorecard process, while 45 percent were dissatisfied. Three statements that managers particularly disagreed with were the scorecard process fairly assessing job performance (48 percent disagreeing versus 31 percent agreeing), the scorecard goals covering all important parts of the job (40 percent disagreeing

---

[20] Survey responses were received from 572 managers, representing a 74 percent response rate.

TABLE 10
1997 Survey Responses by 572 North American Managers on the Balanced Scorecard Process

	Agree	Neutral	Disagree
A. Understanding of the Scorecard Process			
I have a good understanding of the scorecard process	61%	18%	21%
My manager has a good understanding of the scorecard process	71	20	9
With scorecard, it is easy to see the connection between individual and branch performance	41	28	31
B. Implementation and Goal Setting in the Scorecard Process			
The scorecard process is too cumbersome to use effectively	25	31	44
The scorecard goal-setting process is too time consuming to do an adequate job	24	33	43
I am informed about the scorecard goals set for me	68	13	18
I know how my scorecard goals are determined	60	16	24
My scorecard goals can be realistically met, if I work effectively	40	28	33
My scorecard goals cover all the important parts of my job	39	21	40
C. Performance Evaluations and Bonus Awards Using the Scorecard			
The scorecard process fairly assesses my job performance	31	21	48
Bonuses given to people in my unit accurately reflect differences in their performance	30	35	35
When it comes to scorecard bonuses, I have no idea who gets what and why	55	20	25
For employees at my level, bonuses are higher under scorecard than they were under PIP	14	48	39
D. Overall Assessment of the Scorecard Process			
My Area Director believes in the usefulness of the scorecard process	58	38	4
Overall I am satisfied with the scorecard process	32	23	45

versus 39 percent agreeing), and bonuses accurately reflecting differences in their performance (35 percent disagreeing versus 30 percent agreeing).

These attitudes were also reflected in responses given in the open-ended, write-in portion of the survey. The most common issues addressed in this section related to the scorecard being incomplete or overly focused on financial performance (21 percent of the write-in responses). Representative comments included the following:

> I feel we are rated solely on revenues due to the pressure to increase our stock price. No 80 percent employee satisfaction first and 80 percent customer satisfaction second. Revenue is always first and only.
>
> As it stands now, it [expletive deleted]. Scorecard should be more heavily weighted as to service, relationship development, and contributions to the team. "Product pushers" are not the only way to win customers.
>
> Still too much emphasis on "numbers" and not on a "balanced" scorecard.
>
> Our job function entails so much more than the scorecard covers. The scorecard gives a very, very limited picture.

The second most frequent issue related to the inability of managers to understand how their bonuses were determined (14 percent of responses). For example:

> This is a black box process, no one knows anything.
>
> It's too subjective and not objective. I'd prefer an objective rating where everyone concerned knows what to expect when certain levels of performance are achieved.
>
> I would have liked to have known what my evaluation was going to be based on before the quarter began.
>
> I don't understand how it works! I don't see anyone motivated by it. I see the opposite.

Other frequent complaints centered on perceived favoritism, bias, or excessive discretion in scorecard bonus awards (12 percent of responses). Typical comments included the following:

> Eliminate the scorecards! Promotions and bonuses are still given to those who kiss-up to their supervisors, with little regard to performance, educational background, and experience.
>
> I hate the new process. Favoritism comes too much into play.
>
> Some people will get more money, regardless if they do a good job or not, just because they are someone's favorite. Discretion at the AD [area director] level must be eliminated to ensure fairness.

These qualitative analyses indicate that the balanced scorecard at GFS did not achieve its objectives, with GFS experiencing many of the problems identified in behavioral and economic models on the use of subjectivity in reward systems (Prendergast and Topel 1993, 1996). In response, senior management commissioned a team of business unit managers and human resource and compensation staff to review the balanced scorecard plan. The team recommended that the balanced scorecard concept be retained, but that incentive pay be divided into two components: (1) a "structured" or formula-based component, paid quarterly, based on quantitative measures including branch revenues and margins, customer satisfaction indicators, and household growth; and (2) a "discretionary" component, paid annually, based on more qualitative performance objectives such as people leadership, community involvement, projects, and teamwork. By separating bonuses into two distinct segments, the team attempted to retain the benefits of subjectivity while making the link between performance on the quantitative measures and bonus payouts more objective and transparent. Despite this recommendation, the company decided to return to a formula-based bonus plan that rewarded branch managers solely on the basis of branch revenues.

VI. CONCLUSIONS

In this study we examine how different types of performance measures (e.g., financial versus nonfinancial, quantitative versus qualitative, drivers versus results) are weighted in subjective bonus computations. We find that the use of subjectivity in weighting the measures in a balanced scorecard bonus plan allowed supervisors to ignore many performance measures, with financial performance became the primary determinant of bonuses. In addition, the subjectivity in the balanced scorecard plan allowed area directors to incorporate factors other than the scorecard measures in performance evaluations, to change evaluation criteria from quarter to quarter, to ignore measures that were predictive of future financial performance, and to weight measures that were not predictive of desired results. These outcomes led many branch managers to complain about favoritism in bonus awards and uncertainty in the criteria being used to determine rewards, and caused corporate executive and human resource managers to question the scorecard's use for compensation purposes. Ultimately, the balanced scorecard was eliminated in the retail banks, along with many supporting measures such as customer satisfaction and branch quality.

Although it is difficult to generalize results from a single research site, our study provides a number of implications for future research. First, the evidence suggests that psychology-based explanations may be equally or more relevant than economics-based explanations in understanding measurement practices in some settings. In contrast to economic theories, we find little evidence that the weights placed on nonfinancial measures had any association with their ability to predict financial performance (our proxy for informativeness). Instead, we find that most of the weight was placed on quantitative, outcome-oriented financial measures that were used in earlier bonus plans, results that are consistent with psychology-based predictions. This evidence complements prior empirical research indicating that factors other than informativeness influence performance measure choices (e.g., Merchant 1989; Ittner and Larcker 2002).

Second, the study highlights the evolutionary nature of compensation plans. This is an important issue that is missing in pure cross-sectional studies, thereby limiting our understanding of how firms redesign their reward systems in response to unanticipated consequences. Third, our detailed examination of GFS's scorecard plan provides evidence on the multifaceted and complex nature of subjectivity. Prior empirical and theoretical studies have typically examined one element of subjectivity at a time (e.g., flexibility in weighting quantitative performance measures when computing a manager's bonus, the use of qualitative performance evaluations, or the discretion to adjust bonus awards based on factors other than the measures specified in the bonus contract). In contrast, our evidence indicates that all of these elements of subjectivity can occur simultaneously. Further analysis of the use and consequences of subjectivity will need to consider the potential simultaneous presence of these different elements. Finally, our quantitative and qualitative analyses indicate that implementation issues may be far more important to the success or failure of balanced scorecard systems than the scorecard's technical attributes (e.g., the number and types of measures, their classification into categories, or the presence of a causal business model). Consequently, future research on scorecard adoption and performance consequences must move beyond the measurement of these attributes to encompass the entire implementation process.

APPENDIX A
Evolution of the Formula-Based Performance Incentive Plan (PIP) System

Year	Hurdles	Performance Objectives	Bonus for Meeting Performance Targets	Additional Bonus for Exceeding Performance Targets
1993	Satisfaction with primary branch—top 75% of the Western region	Margin growth	3%	—
		Tier I and II household growth	2%	—
		Consumer checking balance growth	2%	—
		B&P checking balance growth	2%	—
		Revenue growth	2%	—
		Liability relationship growth	2%	—
		Expense control	1%	—
		Expenses/revenues	1%	—
1994	Satisfaction with primary branch—statistically at or above the region mean	Margin growth	3%	Two steps, to 1.5%
		Tier I and II household growth	1.5%	Two steps, to 2.5%
	Operations control—audit score of "4" or "5"	Consumer checking balance growth	1.5%	Two steps, to 2.5%
		B&P checking balance growth	1.5%	Two steps, to 2.5%
		Revenue growth	3%	Two steps, to 4.5%
		Liability relationship growth	1.5%	Two steps, to 2.5%
		Expenses/revenues	0.5%	Two steps, to 1%
		Footings/tier I and II households	0.5%	Two steps, to 1%
1995	Branch quality index—at or above the region mean	Overall GFS satisfaction ≥ 80%	5%	—
		Target household growth	2% for growth	Two steps, to 1%
	Operations control—audit score of "4" or "5"	Total checking balance growth	1% for growth 1% for goal	Two steps, to 0.5% Two steps, to 0.5%
		Liability/asset growth	1% for growth 1% for goal	Two steps, to 0.5% Two steps, to 0.5%
	Revenues and margins must meet accountability targets	Revenue growth	2% for growth 2% for goal	Two steps, to 0.5% Two steps, to 0.5%
		Margin growth	2.5% for growth 2.5% for goal	Two steps, to 1% Three steps, to 10%

APPENDIX B
Variable Definitions

Individual Par Score—A rating given to the branch manager for each measure in the balanced scorecard (financial, strategy, customer, control, people, and standards). The par score was expressed as either above par, at par, or below par.

Overall Par Score—A rating given to the branch manager based in part on the six individual par scores for each balanced scorecard category. The par score was expressed as either above par, at par, or below par.

Revenue—Revenue attributed to the branch.

Expense—Expenses matched to branch revenue.

Margin—Revenue minus expenses.

GFS CSI—Score for overall customer satisfaction with overall GFS. This score was expressed as the percentage of survey respondents that rated overall GFS satisfaction in the top two categories on a seven-point scale. A single item was used to measure CSI.

Branch Quality Index—Composite score from survey responses to 20 items. The most heavily weighted item in the branch quality index (45 percent) asked customers to rate "the overall quality of [the branch's] service against your expectations" on a five-point scale. The remaining items include the quality of tellers versus expectations (7.5 percent), six additional items concerning tellers (7.5 percent), quality of other branch personnel versus expectations (7.5 percent), six additional items concerning non-teller employees (7.5 percent), quality of automated teller machines (ATMs) versus expectations (7.5 percent), three additional items concerning ATMs (7.5 percent), and one item measuring problem incidence (10 percent). The composite score was translated into an index with a range from 0 to 100.

Household—A group that makes banking decisions as a family or small business.

Retail households—Individual or family branch customers.

B&P households—Business and professional (B&P) branch customers (essentially small businesses).

Premier households—Customers with at least $100,000 combined balances (e.g., checking, saving, investment, etc., balances). Premier customers have access to a variety of services that are not available to other customers.

Household attrition—Number of households (retail, B&P, or premier) that are no longer branch customers.

Lost households—Same as household attrition.

New households—Number of new groups (retail, B&P, or premier) added to the branch.

Cross sales—Revenue generated by GFS service offerings outside of the branch services (i.e., from branch households).

CNR—Customer Net Revenue or the fees and charges paid by branch customers.

AUM—Assets Under Management by branch customer (including checking, saving, etc. balances).

Automated transactions—Banking services completed using technology such as automated teller machines or the Internet.

Salary—Fixed salary paid to branch manager.

Bonus—Additional incentive payment to branch managers above fixed salary. Bonuses are paid on a quarterly basis.

Labor Grade—An internal GFS human resources category assigned to a branch manager based on individual qualifications. The labor grade is used to set the salary and total compensation ranges (or bands) for an individual branch manager. There are three distinct labor grades used for branch managers.

Time period indicators—Indicator variables that are coded 1 for a specific quarter, and 0 otherwise.

REFERENCES

Baiman, S., and M. V. Rajan. 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review* 70 (4): 557–579.

Baker, G., R. Gibbons, and K. J. Murphy. 1994. Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics* 109 (4): 1125–1156.

Banker, R. D., and S. M. Datar. 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* 27 (1): 21–39.

Banker, R., G. Potter, and D. Srinivasan. 2000. An empirical investigation of an incentive plan that included nonfinancial performance measures. *The Accounting Review* 75 (1): 65–92.

Baron, J., and J. Hershey. 1988. Outcome bias in decision evaluation. *Journal of Personality and Social Psychology* 54 (4): 569–579.

Bellows, R. M. 1954. *Psychology of Personnel in Business and Industry*. Englewood Cliffs, NJ: Prentice Hall.

Bommer, W. H., J. L. Johnson, G. A. Rich, P. M. Podsakoff, and S. B. MacKenzie. 1995. On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology* 48 (3): 587–605.

Blum, M. L., and J. C. Naylor. 1968. *Industrial Psychology*. New York, NY: Harper & Row.

Bushman, R., R. Indjejikian, and A. Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting and Economics* 21 (2): 161–193.

Campbell, D., S. Datar, S. Kulp, and V. G. Narayanan. 2002. Using the balanced scorecard as a control system for monitoring and revising corporate strategy. Working paper, Harvard University.

Campbell, J. P. 1990. Modeling the performance prediction problem in industrial and organizational psychology. In *Handbook of Industrial and Organizational Psychology*, Vol. 1, 2nd edition, edited by M. D. Dunnette, and L. M. Hough, 687–732. Palo Alto, CA: Consulting Psychologists Press.

Datar, S., S. Kulp, and R. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39 (1): 75–92.

Feldman, J. M. 1981. Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology* 66 (1): 127–148.

Feltham, G., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review* 69 (3): 429–453.

Frederickson, J. R., S. A. Pfeffer, and J. Pratt. 1999. Performance evaluation judgments: Effects of prior experience under different performance evaluation schemes and feedback frequencies. *Journal of Accounting Research* 37 (1): 151–165

Ghosh, D., and R. F. Lusch. 2000. Outcome effect, controllability and performance evaluation of managers: Some field evidence from multi-outlet businesses. *Accounting, Organizations and Society* 25 (4–5): 411–425.

Gibbs, M., K. Merchant, W. Van der Stede, and M. Vargus. 2002. Causes and effects of subjectivity in incentives. Working paper, University of Chicago and University of Southern California.

Gjesdal, F. 1981. Accounting for stewardship. *Journal of Accounting Research* 19 (2): 208–231.

Hauser, J. R., D. I. Siemester, and B. Wernerfelt. 1994. Customer satisfaction incentives. *Marketing Science* 13 (3): 327–350.

Hawkins, S. A., and R. Hastie. 1990. Hindsight: Biased judgment of past events after the outcomes are known. *Psychological Bulletin* 107 (3): 311–327.

Hemmer, T. 1996. On the design and choice of "modern" management accounting measures. *Journal of Management Accounting Research* 8 (1): 87–116.

Heneman, R. L. 1986. The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. *Personnel Psychology* 39 (4): 811–826.

———, M. L. Moore, and K. N. Wexley. 1987. Performance-rating accuracy: A critical review. *Journal of Business Research* 15 (5): 431–448.

Holmstrom, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 10: 74–91.

———, and P. Milgrom. 1991. Multi-task principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* 7 (1): 24–52.

Hopwood, A. G. 1974. Leadership climate and the use of accounting data in performance evaluation. *The Accounting Review* 49 (1): 485–495.

Ittner, C., D. Larcker, and M. Rajan. 1997. The choice of performance measures in annual bonus contracts. *The Accounting Review* 72 (2): 231–255.

———, and ———. 2002. Determinants of performance measure choices in worker incentive plans. *Journal of Labor Economics* 20 (2): S58–S90.

Kaplan, R. S., and D. P. Norton. 1996. *The Balanced Scorecard: Translating Strategy into Action.* Boston, MA: Harvard Business School Press.

———, and ———. 2001. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment.* Boston, MA: Harvard Business School Press.

Lambert, R. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32 (1–3): 3–87.

Lipe, M. G. 1993. Analyzing the variance investigation decision: The effects of outcomes, mental accounting, and framing. *The Accounting Review* 68 (4): 748–764.

———, and S. Salterio. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review* 75 (3): 283–298.

———, and ———. 2002. A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society* 27 (6): 531–540.

Luft, J. and M. Shields. 2001. The effects of financial and nonfinancial performance measures on judgment and decision performance. Working paper, Michigan State University.

MacLeod, W. B. 2001. On optimal contracting with subjective evaluation. Working paper, University of Southern California.

Malina, M. A., and F. H. Selto. 2001 Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research* 13 (1): 47–90.

Merchant, K. A. 1989. *Rewarding Results: Motivating Profit Center Managers.* Boston, MA: Harvard University Press.

Mitchell, T., and L. Kalb. 1981. Effects of outcome knowledge and outcome valence on supervisors' evaluations. *Journal of Applied Psychology* 66 (4): 604–612.

Murphy, K., and P. Oyer. 2001. Discretion in executive incentive contracts: Theory and evidence. Working paper, University of Southern California and Stanford University.

Prendergast, C., and R. Topel. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37 (2-3): 355–365.

———, and ———. 1996. Favoritism in organizations. *Journal of Political Economy* 104 (5): 958–978.

Schiff, A. D., and L. R. Hoffman. 1996. An exploration of the use of financial and nonfinancial measures of performance by executives in a service firm. *Behavioral Research in Accounting* 8 (1): 134–151.

Slovic, P., and B. Fischhoff. 1977. On the psychology of experimental surprises. *Journal of Experimental Psychology. Human Perceptions and Performance* 3 (4): 544–551.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 759–778

The Effect of Quality Assessment and Directional Goal Commitment on Auditors' Acceptance of Client-Preferred Accounting Methods

Kathryn Kadous
Emory University

S. Jane Kennedy
University of Washington

Mark E. Peecher
University of Illinois at Urbana–Champaign

ABSTRACT: Previous research demonstrates that auditors' directional goals influence their reporting decisions. For example, when auditors have goals of accepting client-preferred accounting methods, they tend to exploit ambiguity in reporting standards to justify those methods, even when they are aggressive (Hackenbrack and Nelson 1996). We report an experimental investigation of the likely effectiveness of regulation designed to curb this tendency. Specifically, regulators suggest that having auditors identify benchmarks or assess the quality of various methods will "raise the bar" for method acceptability, thereby reducing auditor acceptance of aggressive reporting methods. However, this reasoning ignores the fact that ambiguity typically surrounds quality assessment. Following motivated reasoning theory, we argue that, in order to meet the increased standard for acceptability, auditors with high commitment to directional goals will exploit the ambiguity surrounding the quality of various methods when making quality assessments, with the result that the client-preferred method will be deemed best, or at least of high enough relative quality to be used. This theory suggests that auditor acceptance will increase with goal commitment, and that the increase will be most dramatic when quality assessment is performed. Results of our experiment support our hypotheses that performing a quality assessment *amplifies* the effects of auditors' directional goals on their acceptance of client-preferred methods and on their ratings of the quality of that method. Moreover, auditors making quality assessments

---

We thank V. Arunachalam, S. Biggs, R. Bowen, C. Brown, S. Curtis, E. Hirst, V. Johnson, L. Koonce, R. Libby, M. Lipe, A. Magro, E. Mauldin, L. Myers, M. Nelson, F. Phillips, J. Rich, S. Salterio, L. Sedor, D. Shores, I. Solomon, R. Tubbs, S. Vandervelde, anonymous referees, workshop participants at Cornell University and at the Universities of Connecticut, Illinois at Urbana–Champaign, Iowa, Missouri, Texas at Austin, and Washington, as well as attendees of the 2000 Auditing Section Midyear Conference and the 2000 Behavioral Decision Research in Management Conference for helpful comments.

Editor's note: This paper was accepted by Marlys Gascho Lipe, Editor.

Submitted October 2000
Accepted January 2003

are more likely to identify the client's method as *the most appropriate* method when they are more committed to their directional goals. An implication of our theory and results is that regulation (such as SAS No. 90) that requires auditors to make quality assessments may decrease auditors' objectivity when auditors have directional goals to accept client methods.

Keywords: *objectivity; motivated reasoning; financial reporting; earnings quality.*

Data Availability: *Contact authors.*

I. INTRODUCTION

The SEC has expressed concern that companies are not using the most appropriate accounting methods for their circumstances, and, as a result, the quality of financial reporting is suffering (Levitt 1998). A recent auditing standard, SAS No. 90, *Audit Committee Communications*, addresses this issue by requiring auditors to discuss the quality of client accounting methods with audit committees (AICPA 2000). Some argue that extending auditors' responsibilities beyond compliance to include assessments of the quality or appropriateness of clients' accounting methods will improve auditor objectivity, thereby reducing auditors' willingness to allow clients to use their preferred, aggressive reporting methods (Advisory Panel on Auditor Independence 1994). Ultimately, it is argued, improved auditor objectivity will elevate the quality of financial reporting (Public Oversight Board [POB] 1995; Whitehead et al. 1999). Others argue that unintentional self-serving biases preclude auditors from being objective about client accounting, and so attempts to induce objectivity are unlikely to be successful (Bazerman et al. 1997). The purpose of this paper is to investigate whether requiring auditors to identify the highest quality (i.e., most appropriate) accounting method could actually *decrease* their objectivity by amplifying their tendency to support aggressive, client-preferred accounting methods.

We base our investigation on an extension of motivated reasoning theory. Motivated reasoning theory holds that individuals who are committed to directional goals engage in biased reasoning to reach those goals (Kunda 1990, 1999). Directional goals imply that one conclusion is preferred above others. Individuals with directional goals search for, interpret, and process information in a biased manner and, consequently, are more likely to reach the preferred conclusion (Kunda 1990). Consistent with auditors having a directional goal of supporting client method choices, accounting research demonstrates that auditors tend to interpret ambiguous standards in ways that lend support to client-preferred accounting methods (Hackenbrack and Nelson 1996). We expect that the higher auditors' commitment to such directional goals, the more likely they are to accept client-preferred methods, as long as the methods can be justified (i.e., as long as reasonableness constraints are not binding).

Individuals will not reach preferred conclusions that cannot be justified, regardless of their commitment to a directional goal (Kunda 1990). For example, auditors support client-preferred methods when in-house precedents are ambiguous, but not when precedents unambiguously support alternative treatments (Salterio and Koonce 1997). In general, divergence of the preferred conclusion from salient benchmarks threatens reaching preferred conclusions (Pyszczynski and Greenberg 1987; Kruglanski 1989). Thus, a potential solution to auditors' tendency to support client-preferred methods is to make benchmarks more salient. The Advisory Panel on Auditor Independence (1994) suggests that this can be done by asking auditors to identify the best method or to discuss the relative quality of alternative methods that could be applied. If auditors make objective quality assessments, then the quality gap between the aggressive method that the client prefers and the most appropriate method may cause auditors to disallow the client-preferred method for financial reporting purposes. This approach is essentially the one taken in SAS No. 90 (AICPA 2000).

We predict, however, that auditors will not make quality assessments in an objective manner under some commonly occurring conditions. Extending motivated reasoning theory, we predict that when there is ambiguity surrounding the identity of the best method and the relative quality of alternative methods, auditors will engage in directional processing when asked to assess accounting method quality. Moreover, requiring a quality assessment will threaten auditors' ability to achieve their directional goal by making alternative high-quality methods salient, potentially raising the bar for reporting quality (i.e., instead of determining whether a method is merely acceptable, auditors may consider whether the method is close enough in quality to the best method). As a result, auditors will increase the intensity of effort devoted to justifying the client-preferred method. This additional justification effort will increase both the perceived quality and acceptability of that method. Thus, we expect that performing a quality assessment will amplify the influence of directional goals on auditors' decisions about the acceptability of a client's method.

We test for the presence of this ordinal interaction of quality assessment and directional goal commitment in a Web-based experiment in which we asked auditors to determine whether an aggressive client-preferred accounting method is acceptable. We manipulated quality assessment by asking half the participants to identify the most appropriate accounting method before deciding on the acceptability of the client's method. We measure commitment to the goal of supporting the client method with a validated scale from the organizational behavior literature. We manipulated engagement pressure to induce variability in auditors' directional goals (i.e., to support the use of the client's preferred method) and to address generalizability of our theory. In the lower engagement pressure condition, the auditor learned of the accounting issue after the transaction had been executed but before public reporting. In the higher pressure condition, the auditor learned of the issue only after interim financial statements using the method had been publicly released.

Results support our hypotheses that performing a quality assessment amplifies the effect of directional goal commitment on auditors' acceptance rates and their assessments of the quality of the client-preferred method. In other words, performing a quality assessment *reduces* auditor objectivity. Auditors who make a quality assessment are more likely to accept the client's method when they are more committed to their directional goals to do so, whereas auditors not making a quality assessment are less susceptible to directional goal commitment. Auditors' ratings of the appropriateness of the client-preferred method follow the same pattern. Further, auditors making quality assessments are more likely to identify the client-preferred method as *the most appropriate* method when their commitment to their directional goals is greater. Thus, it appears that instead of improving the quality of financial reporting by making it more difficult for auditors to accept aggressive client-preferred reporting methods, quality assessment causes auditors to ascribe higher quality to these methods, increasing the likelihood that such methods are deemed suitable for financial reporting. An implication of the theory and findings is that regulation requiring auditors to identify benchmarks or assess the quality of methods (such as SAS No. 90) is unlikely to improve, and may even reduce, auditor objectivity when auditors are committed to the directional goal of supporting clients' choices.

The remainder of the paper is organized as follows. The Section II explains the theory and hypotheses. The Sections III and IV describe the experimental method and the results, respectively. The Section V provides discussion and concluding comments.

II. THEORY AND HYPOTHESES

Directional Goals and Auditor Acceptance of Client-Preferred Accounting Methods

The motivated reasoning literature demonstrates that individuals' goals influence their decision processes (e.g., Kunda 1990, 1999). In particular, individuals committed to directional goals, or goals to reach a preferred conclusion, are more likely to reach their desired conclusion, subject to reasonableness constraints (Kunda 1990). Directional reasoning employs a variety of mechanisms. Individuals committed to directional goals conduct biased searches for and overweight the importance of evidence that supports the desired conclusion (Lundgren and Prislin 1998). They choose statistical decision rules and construct new beliefs so as to increase the likelihood that the desired conclusion will be reached (Pyszczynski and Greenberg 1987; Kunda 1990). In addition, they are more skeptical of information that is inconsistent with their preferences than of information that is consistent with their preferences (Ditto and Lopez 1992; Ditto et al. 1998) and they process preference-inconsistent information in a more error-prone manner (Hales 2002). The result of this biased reasoning is that decision makers with higher commitment to directional goals perceive that goal-consistent conclusions are of higher quality, and thus more acceptable, than they otherwise would.

Prior research in accounting is consistent with auditors having directional goals to support client management's method choices and engaging in directional processing. For example, audit seniors tend to choose client-preferred methods and exploit ambiguity in financial accounting standards to justify those methods (Hackenbrack and Nelson 1996). In addition, very experienced auditors tend to make inventory write-down recommendations that are consistent with stated client preferences (Haynes et al. 1998).[1]

Although directional reasoning is pervasive, commitment to directional goals will not always ensure that one will reach goal-consistent decisions. Auditors will reach preferred conclusions only when they can justify them while maintaining the illusion of objectivity (Pyszczynski and Greenberg 1987; Kunda 1990). As a result, directional processing is successful only in conditions of normative ambiguity, or vagueness about what conclusions should or should not be reached. In such cases, one can construct plausible justifications for a variety of conclusions (Pyszczynski and Greenberg 1987). In an auditing context, Johnstone et al. (2001) argue that ambiguity about what alternatives are acceptable is an antecedent to independence impairment. Consistent with this argument, Entwistle and Lindsay (2002) demonstrate that audit partners are more likely to allow clients to use their preferred, aggressive methods when there is greater ambiguity about what methods are acceptable. When normative ambiguity is insufficient, the effect of directional reasoning on decision outcomes is diminished. For example, when in-house precedents unambiguously support alternative treatments (Salterio and Koonce 1997), auditors reduce their tendency to accept aggressive, client-preferred methods.

Effect of Quality Assessment on Directional Reasoning

When a preferred (i.e., goal-consistent) conclusion diverges from salient benchmarks, justifying the preferred conclusion becomes more challenging (Pyszczynski and Greenberg 1987). A sufficiently large divergence between the preferred conclusion and the benchmark threatens the reasonableness of preferred conclusions (Kruglanski 1989), thereby invoking reasonableness constraints. In the financial reporting context, the most appropriate accounting method is likely an important benchmark. Moreover, the availability of a potentially more appropriate method likely threatens the reasonableness of using the client-preferred method for financial reporting purposes (Salterio and Koonce 1997). Thus, a potential solution to the problem of auditors being influenced by directional goals to support client preferences is to make high-quality benchmarks more salient. This could be accomplished by requiring auditors to identify the best method or to assess the relative quality of various methods that could be applied (Advisory Panel on Auditor Independence 1994, 21). SAS No. 90 (AICPA 2000) takes this approach, requiring auditors to make "judgments about the quality, not just the acceptability, of the entity's accounting principles as applied in its financial reporting" and to discuss these judgments with the audit committee.

---

[1] Directional goals arise from a variety of sources. The cited research indicates that auditors' goals tend to be in favor of supporting client management's reporting choices. We expect that these goals arise from social incentives in the form of accountability and from monetary incentives in the form of audit fees or compensation incentives for retaining profitable clients (see Johnstone et al. 2002).

An important question, however, is whether directional goals also bias individuals' judgments about the identity of the benchmark and how favorably preferred conclusions compare with benchmarks. Comparisons to benchmarks are more likely to be effective when the benchmarks are objectively determined and differences between preferred and benchmark conclusions are objectively evaluated. Where there is normative ambiguity surrounding these issues, as is often the case in difficult financial reporting contexts, motivated reasoning theory suggests that benchmarks and relative quality assessments will not be determined in an objective, unbiased manner. Consequently, we expect the influence of an auditor's directional goal to allow the client to use its preferred accounting method to extend to identification of the benchmarks and to evaluations of differences between the preferred method and the benchmark. In other words, auditors will exploit the normative ambiguity surrounding the relative quality of various methods to meet their directional goal of supporting the use of the client-preferred method.

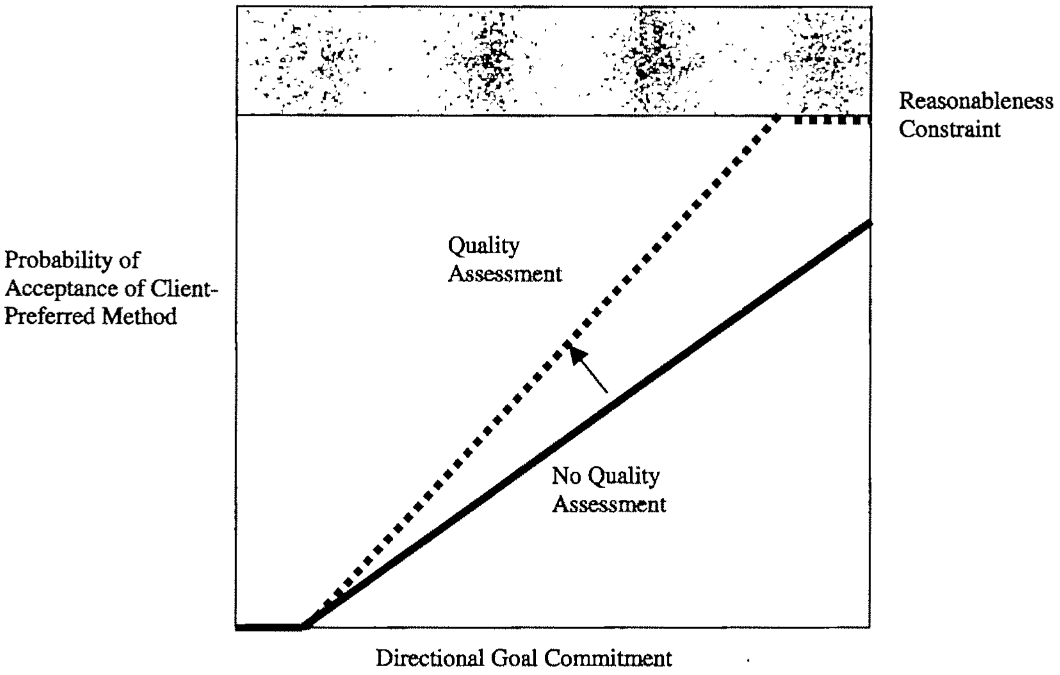
Consistent with this reasoning, Kunda (1990, 487) asserts that asking individuals with directional goals to identify the best conclusion (i.e., a benchmark) may *facilitate* construction of justifications for desired conclusions by increasing the directional processing devoted to reaching those conclusions. Introducing a benchmark likely increases the difficulty of justifying a desired conclusion, so one must work harder to support it. In our financial reporting setting, asking auditors to engage in quality assessment as a precursor to communicating with the audit committee about the overall quality, and not just acceptability, of the client's method raises the possibility that "mere acceptability" may be viewed as an insufficient basis for allowing the client-preferred method for reporting purposes. To be acceptable under such conditions, a method must be the best method or sufficiently close in quality to that method. Thus, when high-quality alternative methods are available, auditors will have to devote effort to directional reasoning to convince themselves, and others, that the client-preferred method is of sufficiently high quality to be used.[2] As long as reasonableness constraints are not binding, as directional goal commitment increases, we expect that performing a quality assessment will amplify auditors' tendency to support the client-preferred method. Accordingly, we expect that quality assessment and directional goal commitment will generate an amplifying form of an ordinal interaction on auditors' propensity to accept client-preferred accounting methods.

Figure 1 illustrates our hypothesized ordinal interaction. The solid line demonstrates the influence of directional goal commitment on auditors' assessments of the quality of the client's method when no quality assessment is performed. When directional goal commitment is very low, directional goals have no discernable influence on acceptance rates.

---

[2] This reasoning is consistent with Boiney et al.'s (1997) finding that the amount of directional processing applied increases only when it becomes apparent that more is needed to ensure attainment of directional goals.

FIGURE 1
Hypothesized Ordinal Interaction of Quality Assessment and Directional Goal Commitment for Acceptance of a Client-Preferred Accounting Method



H_1 predicts that the probability that an auditor will accept a client-preferred method depends jointly on directional goal commitment and whether the auditor is asked to identify the best method (quality assessment). The shaded area represents unreasonable probabilities of acceptance in light of concepts and rules associated with generally accepted accounting principles. At low levels of commitment, directional goals have no discernable influence on the probability of acceptance. However, as directional goal commitment increases, auditors increasingly engage in motivated reasoning and become increasingly likely to accept the client-preferred method (see solid line). Asking auditors to identify a benchmark is expected to cause them to redouble their efforts to meet the directional goal. Thus, quality assessment *amplifies* the effect of directional goal commitment on the probability of acceptance (see dotted line) as long as reasonableness constraints are not binding. A graphical representation of H_2 would be identical except it would involve substituting appropriateness (quality) ratings for probability of acceptance on the y-axis.

However, as directional goal commitment increases, the motivation to ensure use of the client-preferred method becomes stronger. With sufficiently high commitment to directional goals, auditors' reasoning becomes biased in favor of the client-preferred method, increasing the probability of acceptance. Further increases in directional goal commitment continue to increase the probability of acceptance until reasonableness constraints are binding. Because the probability of acceptance is conditional on, instead of independent of, directional goal commitment, auditors can be said to lack objectivity. Performing a quality assessment amplifies the effect of directional goal commitment on auditors' decisions about the acceptability of the client-preferred method, as illustrated by the dotted line. In Figure 1, the steeper positive slope of the dotted line, as compared with that of the solid line, indicates further impairment of auditor objectivity when quality assessment is performed. This pattern of expected acceptance rates is summarized in our first hypothesis:

H_1: Acceptance of the client-preferred method will be highest when quality assessment is performed and goal commitment is relatively high, lower when quality assessment is not performed and goal commitment is relatively high, and lowest when goal commitment is relatively low, regardless of level of quality assessment.

Because we expect this pattern of acceptance to be based on biased assessments of the quality of the client-preferred method, we hypothesize the same ordinal interaction will emerge in ratings of the quality of the client-preferred method. That prediction is our second hypothesis:

H_2: Rated appropriateness of the client-preferred method will be highest when quality assessment is performed and goal commitment is relatively high, lower when quality assessment is not performed and goal commitment is relatively high, and lowest when goal commitment is relatively low, regardless of level of quality assessment.

As noted above, auditors' attempts to meet their directional goals are expected to increase the assessed quality of the client-preferred method. When alternative methods of high quality are available, auditors highly committed to directional goals likely will elevate the judged quality of the client-preferred method to that of the benchmark—the most appropriate method. Our third hypothesis follows:

H_3: Auditors asked to identify the highest quality method will be more likely to identify the client-preferred method as the highest quality method when directional goal commitment is relatively high than when it is relatively low.

III. EXPERIMENTAL METHOD

We test our hypotheses in an experiment in which we manipulated the requirement that the auditor perform a quality assessment (present, absent) between participants and measured commitment to the goal of accepting the client-preferred method (higher, lower). We manipulated engagement pressure (higher, lower) between participants to ensure sufficient variability in the auditors' commitment to directional goals. Participants' primary task was to determine whether to allow the aggressive accounting method proposed by a client in an ambiguous financial reporting situation.

An international accounting firm agreed to let us use its Web-based survey technology to administer our study. Increasingly popular, Web-based research provides access to difficult-to-reach participants who are typically absent from convenience samples (Birnbaum 2001). To date, Web-based findings have been congruent with laboratory findings across diverse research designs, independent variables, and participant pools (Baron and Siepmann 2000; Krantz and Dalal 2000). Further, Web-based data generally are at least as high in quality as laboratory-based data (Birnbaum 1999).

Manipulations and Independent Variables

Quality Assessment

All participants were instructed that they would be asked to decide whether the revenue-recognition method proposed by a new client was acceptable. To accomplish the quality assessment manipulation, we provided half of the participants with additional instructions to identify a benchmark, as follows (italics in original materials):

> Before assessing the *acceptability* of the client-preferred method, however, you will be asked to indicate what you believe to be the *most appropriate* revenue-recognition

> method under GAAP for the circumstances described. The most appropriate method is the one that is most consistent with the conceptual framework and accounting standards and that best captures the economic substance of the transaction under consideration.
>
> As you know, auditors are required to discuss with audit committees the quality, not just the acceptability, of the client company's accounting principles and underlying estimates. The audit committee of the client in this case has a typical level of financial literacy. Your objective assessment of accounting method *appropriateness* will facilitate open and candid discussions with the audit committee about the quality of the client's accounting principles and estimates.

After reading several pages of case materials and immediately before making a decision about the acceptability of the client-preferred method, quality assessment condition participants (only) were asked to identify the most appropriate accounting method. At this point, the instrument included a reminder of the meaning of "most appropriate":

> Recall that the most appropriate method is the one that is most consistent with the conceptual framework and accounting standards and best captures the economic substance of the transaction under consideration.

This language, though not identical to that used in SAS No. 90, captures its focus on the conceptual framework as a benchmark for quality of accounting methods.[3]

Engagement Pressure

We manipulated engagement pressure to ensure variability in auditors' directional goal commitment. Specifically, we varied the timing of the auditor's involvement in the financial reporting issue. In the lower engagement pressure condition, materials indicate that the auditor is performing interim audit testing when the financial reporting issue is identified and that no financial statements containing revenue from the contract has been released to the public. In the higher pressure condition, the materials indicate that the auditor is performing fieldwork testing after year-end and that two sets of interim financial statements using the client-preferred method already have been publicly released.[4]

We constructed the experimental materials such that use of the client-preferred revenue recognition method avoids a negative earnings surprise and keeps intact a prized sequence of annual earnings increases (see Barth et al. 1999), while salient, less aggressive alternative methods do neither. A recent survey suggests that auditors are sensitive to the use of revenue recognition methods in managing earnings (Nelson et al. 2002). To further increase the salience of the public release of information, we included a mock newspaper article describing the first quarter operating results and favorable market reaction to those results in the higher pressure condition. The public release severely limits the client's ability to provide the market with a timely earnings-surprise warning, which could mitigate the client's costs of using an alternative accounting method (Skinner 1994; Kasznik and Lev 1995; Libby and Tan 1999).

---

[3] SAS No. 90 reads, in part "the auditor should discuss with the audit committee the auditor's judgments about the quality, not just the acceptability, of the entity's accounting principles as applied in its financial reporting....The discussion should be tailored to the entity's specific circumstances." (See AICPA 2000.)

[4] To avoid confusion about whether the participant's firm had implicitly approved the method in a timely review, the materials explicitly state in the higher pressure condition that the *predecessor* auditor had reviewed the quarterly statements.

Goal Commitment

Recall that we manipulated engagement pressure to induce variability in commitment to the directional goal of accepting the client-preferred accounting method. Because our hypotheses relate auditors' judgments or decisions to their goal commitment, we use goal commitment rather than engagement pressure as the independent variable in our analyses.

Goal commitment is a measure of motivation and is critical to many organizational behavior theories. Accordingly, researchers in that field have developed and tested scales measuring goal commitment. We used a scale recently refined and validated by Klein et al. (2001), which is based on a longer scale developed by Hollenbeck et al. (1989). The Klein et al. scale requires participants to indicate their level of agreement with five items for each goal on five-point Likert scales. The five items are shown in Exhibit 1. Participants provided agreement ratings for several goals; the goal of interest is to "[b]uild a justifiable case for characterizing [client's] straight-line method to be an acceptable revenue-recognition method in the circumstances."

To develop our measure of goal commitment, we first perform factor analysis on responses to the five items. The eigenvalue for the first factor is 3.11 and all other eigenvalues are less than 1. In addition, the first factor explains 62 percent of the variance in the measures, and Cronbach's alpha is 0.848. This analysis confirms that the five items measure a unitary construct of commitment to the goal of supporting the client-preferred method (e.g., Johnson and Wichern 1988). We perform a median split on this factor score to divide participants into higher and lower goal commitment "conditions" for our subsequent analyses (Johnson and Wichern 1988, 489).

Experimental Case and Procedure

The revenue-recognition issue in our case was adapted from Johnstone et al. (2002). In our version, a first-year audit client has an agreement to provide satellite communication services to a customer at contractual rates, with a predetermined minimum for the first year of the contract. In the lower (higher) pressure condition two (eight) months have passed, and the customer's usage is not keeping pace with the contractual minimum. Johnstone et al. (2002) report that the Big 6 firm with which they developed their case was unable to identify unambiguous professional guidance or precedents for this issue. While we were also unable to identify unambiguous guidance about the most appropriate method, we expected that auditors would view some methods as more appropriate than others for this contract.

EXHIBIT 1
Goal Commitment Measures

1. I thought this was a good goal to shoot for.
2. I was strongly committed to pursuing this goal.
3. It was hard to take this goal seriously. (R)
4. Quite frankly, I didn't care if I achieved this goal or not. (R)
5. It wouldn't have taken much to make me abandon this goal. (R)

We used these five items to measure goal commitment. The items comprise a scale developed by Klein et al. (2001), and were applied to the goal to "[b]uild a justifiable case for characterizing [client's] straight-line method to be an acceptable revenue-recognition method in the circumstances." Items followed by (R) are reverse-scored.

Materials include brief instructions, client background information, a description of the contract, and a description of the client's preferred revenue-recognition method. The client prefers to allocate the minimum-guaranteed contract revenue evenly across the 12 months of the initial contract period. The first eight months of the initial contract period occur in the current fiscal year. The client justifies its straight-line method in terms of simplicity and argues that alternative methods would push significant revenue into the first quarter of next year, generating unrealistic earnings expectations for the remainder of next year. No alternative methods are described in the initial case materials. The client-preferred method is aggressive in that methods based on actual use of satellite time allocate less than half as much revenue to the current year. The difference between revenue under the client-preferred method and methods based on actual satellite use is approximately 3 percent of current-year revenue.

After reading background and contract information, participants in the quality assessment conditions were asked to identify the most appropriate accounting method by either selecting the client-preferred method or identifying a different method. Next, all participants indicated (by binary choice) whether the client's accounting method was acceptable, provided goal commitment ratings, and rated the appropriateness of the client-preferred method (as well as two alternative methods) on a scale from 1 = Not at all Appropriate to 7 = Completely Appropriate.[5] A post-experimental questionnaire included manipulation checks and demographic items. Main dependent measures are participants' decisions about the acceptability of the client-preferred accounting method, assessments of the most appropriate accounting method, and ratings of the appropriateness of the client-preferred method.

IV. RESULTS

Response Rate

The survey administrator announced the study via email to approximately 3,500 of the firm's auditors in the United States with a rank of supervising senior, manager, or partner. Auditors could click on a hyperlink in the email to be connected with one of four Web sites, corresponding to the four conditions in the study. The initial email messages counterbalanced auditor rank and geographic location across conditions. A total of 227 auditors completed the instrument online during a one-week response period during December 2001, yielding an overall response rate of 6.48 percent. This rate, while low, is consistent with the survey administrator's expectations, given the nature and length of the experimental instruments.[6] Mean (standard deviation of) audit experience of respondents was 9.63 (7.26) years.[7]

---

[5] We chose a binary dependent measure because, ultimately, auditors must decide whether to accept or reject client accounting methods. Although continuous response modes also are well suited for testing our theory, it is possible that statistically different judgments may fall on the same side of decision thresholds (e.g., Gilliland et al. 1998). To avoid the risk of identifying differences in judgment that do not result in differences in ultimate decisions, we use this binary measure in our primary analyses.

[6] Because they allow for particularly easy and anonymous withdrawal, Web-based experiments are better than laboratory experiments at detecting participants' aversion to difficult or contentious experimental conditions (Reips 2000). Consistent with differential aversion across conditions, our manipulations of quality assessment and engagement pressure interactively affected the number of responses received (Mantel-Haenszel $\chi^2 = 4.24$, $p = 0.04$). We received the fewest responses in the quality assessment/higher pressure condition ($n = 43$ or 4.9 percent), the next fewest in the no quality assessment/higher pressure condition ($n = 55$ or 6.3 percent), the next fewest in the quality assessment/lower pressure condition ($n = 62$ or 7.1 percent), and the most responses in the no quality assessment/lower pressure condition ($n = 67$ or 7.7 percent).

[7] Mean audit experience is likely understated because the instrument truncated experience at 25 years, and 17 participants recorded 25 or more years of experience.

Manipulation Checks

Quality Assessment

Of the 98 quality assessment condition participants, 28 indicated that the client's method was the most appropriate method and 70 participants indicated that another method was the most appropriate. Sixty-eight of 70 participants who chose an alternative method (97.1 percent) provided the requested written justification for their choice. These results indicate that quality assessment condition participants attended to the quality assessment instruction.[8]

Engagement Pressure

We asked participants in the post-test, "In the case that you read, to what extent was there pressure to accept [the client's] straight-line method?" The response scale had endpoints labeled 1 = Very Low Pressure and 7 = Very High Pressure and a midpoint labeled 4 = Average Pressure. Perceived pressure was higher in the higher pressure condition than in the lower pressure condition (means [standard deviations] are 6.38 [0.98] and 6.01 [1.11], respectively, $t_{161} = 2.21$, one-sided $p = 0.01$), indicating a successful manipulation.[9] Perceived pressure is not associated with quality assessment ($t_{161} = 0.74$, $p = 0.46$), indicating independence of the two manipulations.[10]

Directional Goal Commitment

Goal commitment scores exhibit substantial variation within and across engagement pressure conditions. Given lower engagement pressure, the average (median) goal commitment score is 0.14 (0.20), and the standard deviation (inter-quartile range) is 0.95 (−0.41 to 0.82). Given higher engagement pressure, the average (median) goal commitment score is −0.16 (−0.19), and the standard deviation (inter-quartile range) is 1.03 (−0.95 to 0.80). Analysis of variance (ANOVA) for goal commitment factor scores reveals that average reported goal commitment is significantly *lower* in the higher engagement pressure condition than in the lower engagement pressure condition ($F_{1,223} = 4.86$; $p = 0.03$). Neither the presence or absence of quality assessment ($F_{1,223} = 0.96$, $p = 0.33$) nor the interaction of engagement pressure and quality assessment ($F_{1,223} = 0.12$, $p = 0.73$) influence goal commitment factor scores.

We use a median split on goal commitment factor scores in our analyses, although results are similar when using the continuous goal commitment scores. Mean (standard deviation) scores for the high and low commitment conditions are 0.81 (0.58) and −0.81 (0.59), respectively. These values are statistically significantly different from one another ($t_{225} = 10.84$, $p < 0.01$). In the analyses, we report tests for our hypothesized ordinal interaction between directional goal commitment and quality assessment both aggregated across engagement pressure conditions and separately within each of the lower and the higher engagement pressure conditions. Separate tests provide evidence about the generalizability of our theory over levels of engagement pressure.

---

[8] It is possible that participants in the no quality assessment condition spontaneously considered what method would be most appropriate. To the extent that they did so, our tests would be biased against finding effects of the quality assessment manipulation on auditors' decisions and judgments.

[9] All p-values are two-sided except when otherwise noted. We use one-sided p-values only for tests using non-categorical dependent measures for which our theory implies a directional effect.

[10] We also asked participants to identify whether the client had publicly recognized revenue from the new contract in any quarterly financial statements filed with the SEC. Participants in the higher pressure condition were more likely to respond positively than were participants in the lower pressure condition (95.2 versus 15.6 percent; $\chi_1^2 = 143.59$, $p < 0.01$). The proportion of positive responses is independent of quality assessment condition ($\chi_1^2 = 0.14$, $p = 0.71$).

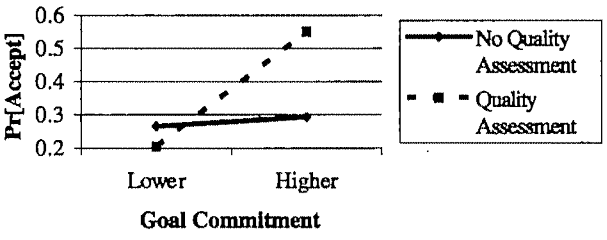
Tests of Hypotheses about Acceptance of the Client-Preferred Method

Mean rates of acceptance of the client-preferred method are graphed in Panel A of Figure 2 for the conditions relevant to our hypotheses. Table 1 provides the percentage of participants who accepted the client's method by conditions—both aggregated over engagement pressure conditions and separately for higher and lower engagement pressure. Overall, 32.2 percent of participants accepted the client's method. We interpret the low acceptance rate as evidence that auditors found justifying the client's method to be challenging.

Our first hypothesis examines whether performing a quality assessment amplifies the influence of auditors' directional goal commitment on their acceptance of the client-preferred method. It defines a specific pattern of acceptance rates across the four experimental conditions that obtain by crossing quality assessment (absent versus present) and the median split on goal commitment factor scores (lower versus higher). Specifically, it predicts that acceptance of the client-preferred, aggressive method is highest when a quality assessment is performed and goal commitment is higher, lower when there is no quality assessment and goal commitment is higher, and lowest in the two lower goal commitment conditions. Accordingly, we test H_1 with a linear contrast of cell means (Buckless and Ravenscroft 1990; Keppel 1991). We use contrast weights of +3 for the quality assessment, higher goal commitment condition, +1 for the no quality assessment, higher goal

FIGURE 2
Observed Acceptance and Appropriateness of the Client-Preferred Accounting Method

Panel A: Acceptance of Client Method (H1)



Panel B: Appropriateness of Client Method (H2)

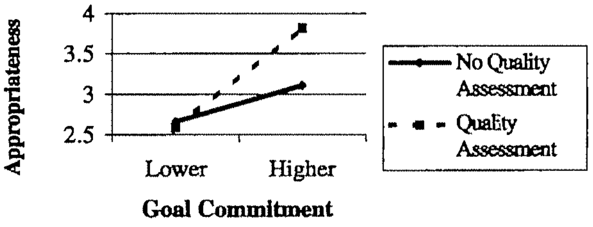


TABLE 1
Proportion (Percentage) of Auditors Accepting the Client-Preferred Method by Condition

	Engagement Pressure (EP)				**Collapsed across Engagement Pressure**		**Collapsed across EP and Goal Commitment**
	Lower		**Higher**				
	Goal Commitment		**Goal Commitment**		**Goal Commitment**		
	Lower	**Higher**	**Lower**	**Higher**	**Lower**	**Higher**	
Quality Assessment	4/21 (19.0)	16/34 (47.1)	6/28 (21.4)	11/15 (73.3)	10/49 (20.4)	27/49 (55.1)	37/98 (37.8)
No Quality Assessment	10/32 (31.3)	11/35 (31.4)	7/33 (21.2)	8/29 (27.6)	17/65 (26.2)	19/64 (29.7)	36/129 (27.9)
Collapsed across Quality Assessment	14/53 (26.4)	27/69 (39.1)	13/61 (21.3)	19/44 (43.2)	27/114 (23.7)	46/113 (40.7)	73/227 (32.2)

Lower (higher) goal commitment refers to below- (above-) median factor scores based on participants' responses to the Klein et al. (2001) scale items (see Exhibit 1) measuring commitment to the goal to "[b]uild a justifiable case for characterizing [client's] straight-line method to be an acceptable revenue-recognition method in the circumstances."

Quality assessment was manipulated by asking auditors to identify the most appropriate method for the client's situation before asking whether the client-preferred method is acceptable (quality assessment condition), or not asking auditors to identify the most appropriate method (no quality assessment condition).

Engagement pressure was manipulated by varying the timing of auditor involvement in the financial reporting issue. In the lower engagement pressure condition, the auditor became aware of the accounting issue before financial information using the client-preferred method was released. In the higher engagement pressure condition, the auditor became aware of the issue after two sets of interim statements using the client-preferred method had been released.

The dependent measure reflects participants' (binary) decisions as to whether the client-preferred, aggressive accounting method is acceptable.

commitment condition and –2 for each of the lower goal commitment conditions (Rosnow and Rosenthal 1995).[11]

A traditional categorical ANOVA model is provided in Panel A of Table 2, and our hypothesis test (using the above contrast weights) as well as the simple effects of directional goal commitment are provided in Panel B. As reported, the hypothesis test is statistically significant ($\chi_1^2 = 12.27$, $p < 0.01$), supporting H_1. Simple effect tests (also in Panel B) reveal that the influence of directional goal commitment on acceptance rates is not statistically significant when quality assessment is not performed ($\chi_1^2 = 0.11$, $p = 0.74$), but that higher directional goal commitment is associated with higher acceptance of the client-preferred method when quality assessment is performed ($\chi_1^2 = 11.78$, $p < 0.01$).[12]

To test the generalizability of these findings, we examine whether the hypothesized ordinal interaction between quality assessment and goal commitment obtains in both engagement-pressure conditions using a median split on goal commitment within each of

---

[11] Rosnow and Rosenthal (1995, 7) identify these weights as appropriate for the form of our hypothesized ordinal interaction. We do not rely only on the interaction term from the traditional ANOVA model because our hypothesized interaction is ordinal in form, and traditional ANOVA models allocate much of the variance attributable to ordinal interactions to simple main effects instead of the interaction term (Buckless and Ravenscroft 1990; Dawes 1988; Rosnow and Rosenthal 1995). We present the traditional ANOVA results for completeness.

[12] Viewed another way, the simple effects of quality assessment are statistically significant given higher goal commitment ($\chi_1^2 = 7.56$, $p < 0.01$), but not given lower goal commitment ($\chi_1^2 = 0.58$, $p = 0.45$).

TABLE 2
Inferential Statistics Regarding Auditor Acceptance of the Client-Preferred Method

Panel A: Categorical ANOVA

Source	**df**	χ^2	**Two-Sided p-value**
Intercept	1	25.78	<0.01
Quality Assessment	1	1.53	0.22
Goal Commitment	1	7.96	<0.01
Quality Assessment by Goal Commitment	1	5.66	0.02

Panel B: Planned Contrasts

Hypothesized Contrast	**df**	χ^2	**Two-Sided p-value**
H_1: Acceptance is highest in the quality assessment/higher goal commitment condition, lower in the no quality assessment/higher goal commitment condition, and lowest in the two lower goal commitment conditions (contrast weights are +3, +1, –2, and –2, respectively).	1	12.27	< 0.01

Simple Main Effects of Goal Commitment	**df**	χ^2	**Two-sided p-value**
Acceptance is higher for higher versus lower goal commitment for no quality assessment participants.	1	0.11	0.74
Acceptance is higher for higher versus lower goal commitment for quality assessment participants.	1	11.78	< 0.01

See Table 1 for descriptions of all variables.

the lower and higher engagement pressure conditions.[13] In the lower pressure conditions, the hypothesized ordinal interaction is statistically significant ($\chi_1^2 = 4.70$, $p = 0.03$). The simple effect of directional goal commitment is statistically significant when quality assessment is present ($\chi_1^2 = 5.01$, $p = 0.03$), but not when quality assessment is absent ($\chi_1^2 = 0.02$, $p = 0.88$). Given higher pressure, the hypothesized ordinal interaction is again statistically significant ($\chi_1^2 = 12.97$, $p < 0.01$). The simple effect of directional goal commitment is statistically significant when a quality assessment is performed ($\chi_1^2 = 10.74$, $p < 0.01$), but not when quality assessment is absent ($\chi_1^2 = 0.55$, $p < 0.46$).[14]

---

[13] Because these median splits occur within pressure conditions, they occur at different values than the median splits used to test the overall hypothesis (those upon which Tables 1 and 3 are based). In both cases, the average goal commitment factor score is significantly lower in the higher engagement pressure condition than in the lower engagement pressure condition.

[14] If we substitute manipulated engagement pressure for the measured goal commitment variable, quality assessment is marginally associated with acceptance ($p = 0.10$), but no other effects are statistically significant. In the presence (absence) of quality assessment, 37.8 percent (27.9 percent) of auditors accept the client method. When only higher goal commitment participants are considered, though, an interaction of quality assessment and engagement pressure, analogous to H_1, is marginally statistically significant ($p = 0.06$).

(continued on next page)

Consistent with H_1, auditors who assess quality appear to be more susceptible to their own directional goals (i.e., are less objective) in making decisions about whether to accept a client-preferred accounting method. Even though auditors in higher pressure conditions reported lower commitment to the pro-client directional goal, the interaction between quality assessment and goal commitment is not weaker with higher engagement pressure.

Tests of Hypotheses about Appropriateness of the Client-Preferred Method

Cell means of rated appropriateness, or quality, of the client-preferred, aggressive accounting method are graphed in Figure 2, Panel B. Cell means and standard deviations, both aggregated across engagement pressure levels and separately by engagement pressure level, are provided in Table 3.

In H_2, we predict that rated appropriateness of the client-preferred accounting method will increase with goal commitment, and that the increase will be larger for auditors in the quality assessment condition. In other words, we expect appropriateness ratings to follow the same pattern as hypothesized for acceptance decisions. Accordingly, we use the same contrast weights to test H_2 that we used to test H_1. A traditional ANOVA analysis and the applicable contrast tests appear in Panels A and B of Table 4, respectively.

As indicated in Panel B of Table 4, the hypothesized contrast is statistically significant ($t_{223} = 3.64$, one-sided $p < 0.01$), supporting H_2. Simple effects tests show that directional goal commitment is statistically significant when a quality assessment is performed ($t_{223} = 3.22$, one-sided $p < 0.01$), but not when quality assessment is not performed ($t_{223} = 1.27$,

TABLE 3
Mean (Standard Deviation) Auditor Ratings of the Appropriateness of the Client-Preferred Method by Condition

	Engagement Pressure (EP)				Collapsed across Engagement Pressure		
	Lower		Higher				
	Goal Commitment		Goal Commitment		Goal Commitment		Collapsed across EP and Goal Commitment
	Lower	Higher	Lower	Higher	Lower	Higher	
Quality Assessment	2.47	3.61	2.67	4.27	2.59	3.82	3.20
	(1.89)	(2.03)	(2.07)	(2.12)	(1.98)	(2.06)	(2.10)
No Quality Assessment	3.00	2.75	2.58	3.24	2.66	3.11	2.88
	(1.65)	(1.81)	(2.04)	(1.57)	(1.92)	(1.60)	(1.78)
Collapsed across Quality Assessment	2.64	3.30	2.62	3.59	2.63	3.41	3.02
	(1.83)	(1.86)	(2.04)	(1.82)	(1.94)	(1.84)	(1.93)

The dependent measure reflects participants' evaluations, on a scale ranging from 1 to 7, of the appropriateness of the client-preferred, aggressive accounting method.
See Table 1 for descriptions of the independent variables.

Footnote 14, continued

We conducted an earlier pencil-and-paper experiment that was very similar to the Web-based experiment reported herein, except that participants predominantly were audit senior attendees at a national training session for new managers, and no participants dropped out from the experiment. We did not measure goal commitment in that experiment. In that experiment, an interaction of quality assessment and (manipulated) engagement pressure, analogous to H_1, obtained ($p = 0.01$).

TABLE 4

Inferential Statistics Regarding Rated Appropriateness of the Client-Preferred Method

Panel A: ANOVA

Source	**df**	**MS**	**F**	**Two-Sided p-value**
Quality Assessment	1	5.77	1.63	0.20
Goal Commitment	1	37.67	10.62	<0.01
Quality Assessment by Goal Commitment	1	9.00	2.54	0.11
Residual	223	3.55		

Panel B: Planned Contrasts

Hypothesized Contrast	**df**	**t**	**One-Sided p-value**
H_2: Rated appropriateness of the client-preferred method is highest in the quality assessment/higher goal commitment condition, lower in the no quality assessment/higher goal commitment condition, and lowest in the two lower goal commitment conditions (contrast weights are +3, +1, −2, and −2, respectively).	223	3.64	<0.01
Simple Main Effects of Goal Commitment	**df**	**t**	**One-Sided p-value**
Rated appropriateness of the client-preferred method is higher for higher versus lower goal commitment for no quality assessment participants.	223	1.27	0.10
Rated appropriateness of the client-preferred method is higher for higher versus lower goal commitment for quality assessment participants.	223	3.22	<0.01

See Table 1 for descriptions of the independent variables.
See Table 3 for a description of the dependent variable.

one-sided $p = 0.10$).[15] As we did for acceptance rates, we perform separate tests for the hypothesized ordinal interaction in the two engagement pressure conditions using median splits on goal commitment within each pressure condition. In the lower pressure condition, the hypothesized ordinal interaction is statistically significant ($t_{118} = 2.66$, one-sided $p < 0.01$). The simple effect of directional goal commitment is statistically significant when quality assessment is present ($t_{118} = 2.59$, one-sided $p < 0.01$), but not when quality assessment is absent ($t_{118} = 0.50$, one-sided $p = 0.31$). When pressure is higher, the hypothesized ordinal interaction is again statistically significant ($t_{101} = 3.28$, one-sided $p < 0.01$). The simple effect of directional goal commitment is statistically significant both when a quality assessment is performed ($t_{101} = 2.71$, one-sided $p < 0.01$) and when a quality assessment is not performed (t_{118} = 1.66, one-sided p = 0.05).[16] This pattern of results, observed both for quality ratings for the client-preferred method (H_2) and acceptance decisions about that method (H_1), suggests that requiring auditors with directional goals to perform quality assessments can intensify directional processing.[17]

---

[15] The simple effects of quality assessment are statistically significant given higher goal commitment ($t_{223} = 2.03$, one-sided $p = 0.02$) but not given lower goal commitment ($t_{223} = 0.22$, $p = 0.82$)

Tests of Hypothesis about Identification of Client's Method as Most Appropriate

H_3 is also supported. For participants making a quality assessment, 19 of 49 (39 percent) participants with higher commitment to the directional goal identified the client-preferred method as most appropriate, while only nine of 49 (18 percent) lower goal commitment participants did so. This difference is statistically significant (Fisher's Exact Test, p = 0.02). Consistent with our expectations, auditors performing a quality assessment are more likely to identify the client-preferred method as the most appropriate method. Taken together, the results for H_2 and H_3 suggest that biased benchmarks and biased ratings of method quality are likely responsible for the joint influence of quality assessment and directional goals on acceptance of the client-preferred method.

V. CONCLUDING REMARKS

Motivated reasoning theory predicts and previous accounting studies suggest that auditors' decisions about the acceptability of client-preferred accounting methods are influenced by auditors' directional goals to accept client-preferred methods (e.g., Hackenbrack and Nelson 1996). In this study, we extend motivated reasoning theory to predict that asking auditors to identify the most appropriate, or benchmark, method will amplify the influence of their directional goals on their acceptance decisions. This issue is relevant to accounting practice because requiring some form of quality assessment has been suggested as a potential way of improving the quality of reporting (e.g., Advisory Panel on Auditor Independence 1994) and is an integral part of a recent auditing standard, SAS No. 90 (AICPA 2000).

Results of our experiment support our hypotheses. In particular, we find that asking auditors to identify the best method increases the influence of commitment to their own directional goals on their tendency to accept a client-preferred method, even though that method is aggressive and methods of potentially higher quality are available. Moreover, auditors are more likely to identify the client-preferred method as *the most appropriate* method when they are more committed to their directional goal.

This study has implications for the audit standard-setting process. SAS No. 90 requires discussion between auditors and audit committees about the relative quality of a client's chosen accounting methods (AICPA 2000). Conventional wisdom suggests that requiring auditors to assess the quality of potential accounting methods will improve auditor objectivity (Advisory Panel on Auditor Independence 1994) and reduce aggressive reporting by providing audit committee members with objective information about the quality of alternative accounting methods (Whitehead et al. 1999). Our study, however, suggests that this standard is likely to *decrease* auditor objectivity when auditors have directional goals to support the client-preferred method. As auditors' commitment to pro-client directional goals increases, requiring quality assessments evidently increases auditors' motivation to engage in reasoning that supports acceptance of a client-preferred method. Quality assessment amplifies the effects of directional goal commitment on auditors' judgment processes, causing them to ascribe higher quality to the client's method than they otherwise would. The result is that benchmarks and quality assessments, and thus acceptance decisions, are *more* biased in favor of the client-preferred method.

---

[16] If we substitute manipulated engagement pressure for the measured goal commitment variable, no effects are statistically significant either overall or when examining lower commitment participants only (all p > 0.22). Among higher goal commitment participants, though, the ordinal interaction between quality assessment and engagement pressure is statistically significant (p = 0.05). In the pencil-and-paper experiment described in footnote 14, we observed an analogous interaction (p = 0.04).

[17] Audit experience has a small, marginally statistically significant, negative correlation with acceptability and appropriateness of the client-preferred method (r = −0.13, p = 0.06 for acceptability and r = −0.12, p = 0.07 for appropriateness). Audit experience is not statistically significantly related to whether the participant identifies the client-preferred method as most appropriate (χ_1^2 = 0.50, p = 0.45). We repeated all analyses while controlling for audit experience, and inferences are identical to those reported. Further, experience does not significantly interact with either quality assessment or goal commitment.

While this study demonstrates the potential dangers of mandating quality assessment in an environment in which auditors have directional goals to support client-preferred methods, it does not thoroughly investigate the conditions under which auditors become more or less committed to pro-client directional goals. It should be noted that our theory predicts that performing a quality assessment amplifies *prevailing* directional goals, regardless of whether these goals favor client management. At present, auditors have significant social and financial incentives to please management (Bazerman et al. 1997; Johnstone et al. 2002), and audit committees are unlikely to develop and effectively communicate reporting preferences that differ from methods preferred by management and supported by the auditor.[18] Consequently, pro-client directional goals generally prevail. However, if the costs of being strongly committed to pro-client directional goals were to increase sufficiently (e.g., due to litigation risk as in Farmer et al. [1987]) or if knowledgeable audit committees were able to hold auditors accountable for high-quality, transparent reports, auditors' directional goals could change. In either case, we would expect quality assessment to amplify the new directional goal.[19]

This study also provides insight into likely impacts of the SEC's recent rule requiring timely reviews of quarterly financial statements on acceptance of client-preferred accounting methods (SEC 2000). We find that for auditors involved in the decision process only after quarterly financial statements have been issued, engagement pressure is higher, but reported goal commitment is lower. Despite lower reported overall levels of goal commitment, our hypotheses are supported in the higher pressure conditions. In fact, results appear strongest in these conditions (e.g., see Tables 1 and 3) and mean acceptance rates and appropriateness ratings for the client-preferred method are ***higher*** when engagement pressure is higher (see Table 2, Panel A for acceptance and Table 4, Panel A for appropriateness). Thus, bias in auditors' quality assessments and acceptance decisions is likely exacerbated by higher pressure. Accordingly, if timely reviews reduce engagement pressure by increasing the likelihood of early auditor involvement in controversial reporting decisions, then the SEC's rule will likely be effective in mitigating aggressive reporting.

This experimental investigation, like others, is limited in the number of real-world features of audit practice that it encompasses. It is possible that including such factors could moderate the effects we observed. For example, we did not collect any measures related to participants' task-specific knowledge. Johnstone et al. (2002) find that knowledge influences the generation of alternative accounting methods for a similar contract. Such knowledge may moderate the extent to which quality assessment and goal commitment interactively affect auditors' propensity to accept, or ascribe high quality to, a client-preferred method.

---

[18] According to a recent survey, 31 percent of audit committees for large firms do not satisfy recommendations regarding financial literacy and independence from management and 52 percent do not meet as frequently as financial statements are issued (Read and Raghunandan 2001).

[19] In a related experiment, we placed all participants (auditing students) under higher engagement pressure and held them accountable to an audit committee for their decisions. We manipulated audit committee member preferences and quality assessment. Half of the audit committees emphasized the costs associated with restatements, while the others emphasized the cost of litigation due to overly aggressive accounting. Consistent with theory, we find that audit committee preferences influence participants' willingness to accept the client-preferred method and that performing quality assessment amplifies the effect of that influence ($\chi_2^2 = 6.61$, $p = 0.04$).

REFERENCES

Advisory Panel on Auditor Independence. 1994. *Strengthening the Professionalism of the Independent Auditor*. Stamford, CT: Public Oversight Board.

American Institute of Certified Public Accountants (AICPA). 2000. *Audit Committee Communications*. Statement on Auditing Standards No. 90. New York, NY: AICPA.

Baron, J., and M. Siepmann. 2000. Techniques for creating and using Web questionnaires in research and teaching. In *Psychological Experiments on the Internet*, edited by M. H. Birnbaum, 235–265. San Diego, CA: Academic Press.

Barth, M. E., J. A. Elliott, and M. W. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37 (2): 387–414.

Bazerman, M., K. Morgan, and G. Loewenstein. 1997. The impossibility of auditor independence. *Sloan Management Review* 38: 89–94.

Birnbaum, M. H. 1999. Testing critical properties of decision making on the Web. *Psychological Science* 10: 399–407.

———. 2001. *Introduction to Behavioral Research on the Internet*. Upper Saddle River, NJ: Prentice Hall.

Boiney, L. G., J. Kennedy, and P. Nye. 1997. Instrumental bias in motivated reasoning: More when more is needed. *Organizational Behavior and Human Decision Processes* 1: 1–24.

Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65 (4): 933–945.

Dawes, R. M. 1988. *Rational Choice in an Uncertain World*. Orlando, FL: Harcourt Brace Jovanovich.

Ditto, P. H., and D. F. Lopez. 1992. Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology* 63 (4): 568–584.

———, J. A. Scepansky, G. D. Munro, A. M. Apanovitch, and L. K. Lockhart. 1998. Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology* 75 (1): 53–69.

Entwistle, G., and D. Lindsay. 2002. Audit partner judgment: The effects of concurring reviews, compensation practices, and accounting flexibility. Working paper, University of Saskatchewan.

Farmer, T. A., L. E. Rittenberg, and G. M. Trompeter. 1987. An investigation of the impact of economic and organizational factors on auditor independence. *Auditing: A Journal of Practice & Theory* 7 (1): 1–14.

Gilliland, S. W., L. Benson, and D. H. Schepers. 1998. A rejection threshold in justice evaluation: Effects on judgment and decision-making. *Organizational Behavior and Human Decision Processes* 76 (2): 113–131.

Hackenbrack, K., and M. W. Nelson. 1996. Auditors' incentives and their application of financial accounting standards. *The Accounting Review* 71 (1): 43–59.

Hales, J. W. 2002. Understanding bias and dispersion in forecasts: The role of motivated reasoning. Working paper, Cornell University.

Haynes, C. M., J. G. Jenkins, and S. R. Nutt. 1998. The relationship between client advocacy and audit experience: An exploratory analysis. *Auditing: A Journal of Practice & Theory* 17 (2): 88–104.

Hollenbeck, J. R., C. R. Williams, and H. J. Klein. 1989. An empirical examination of the antecedents of commitment to difficult goals. *Journal of Applied Psychology* 74 (1): 18–23.

Johnson, R. A., and D. W. Wichern. 1988. *Applied Multivariate Statistical Analysis: Second Edition.* Englewood Cliffs, NJ: Prentice Hall.

Johnstone, K. M., M. H. Sutton, and T. D. Warfield. 2001. Antecedents and consequences of independence risk: Framework for analysis. *Accounting Horizons* 15 (1): 1–18.

———, J. C. Bedard, and S. F. Biggs. 2002. Aggressive client reporting: Factors affecting auditors' generation of financial reporting alternatives. *Auditing: A Journal of Practice & Theory* 21 (1): 47–65.

Kasznik, R., and B. Lev. 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review* 70 (1): 113–134.

Klein, H. J., M. J. Wesson, J. R. Hollenbeck, P. M. Wright, and R. P. DeShon. 2001. The assessment of goal commitment: A measurement model meta-analysis. *Organizational Behavior and Human Decision Processes* 85 (1): 32–55.

Keppel, G. 1991. *Design and Analysis: A Researcher's Handbook.* Upper Saddle River, NJ: Prentice Hall.

Krantz, J. H., and R. Dalal. 2000. Validity of Web-based psychological research. In *Psychological Experiments on the Internet,* edited by M. H. Birnbaum, 35–60. San Diego, CA: Academic Press.

Kruglanski, A. W. 1989. The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin* 106 (3): 395–409.

Kunda, Z. 1990. The case for motivated reasoning. *Psychological Bulletin* 108: 480–498.

———. 1999. *Social Cognition: Making Sense of People.* Cambridge, MA: The MIT Press.

Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, New York, NY, September 28.

Libby, R., and H. Tan. 1999. Analysts' reactions to warnings of negative earnings surprises. *Journal of Accounting Research* 37 (2): 415–435.

Lundgren, S. R., and R. Prislin. 1998. Motivated cognitive processing and attitude change. *Personality and Social Psychology Bulletin* 24 (7): 715–726.

Nelson, M., J. Elliott, and R. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.

Public Oversight Board (POB). 1995. *Directors, Management, and Auditors: Allies in Protecting Shareholder Interests.* Stamford, CT: POB.

Pyszczynski, T., and J. Greenberg. 1987. Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. *Advances in Experimental Social Psychology* 20: 297–340.

Read, W. J., and K. Raghunandan. 2001. The state of audit committees. *Journal of Accountancy* (May): 57–60.

Reips, U. 2000. The Web experiment method: Advantages, disadvantages, and solutions. In *Psychological Experiments on the Internet,* edited by M. H. Birnbaum, 89–117. San Diego, CA: Academic Press.

Rosnow, R. L., and R. Rosenthal. 1995. "Some things you learn aren't so": Cohen's paradox, Asch's paradigm, and the interpretation of interaction. *Psychological Science* 6 (1): 3–9.

Salterio, S., and L. Koonce. 1997. The persuasiveness of audit evidence: The case of accounting policy decisions. *Accounting Organizations and Society* 22: 573–587.

Securities and Exchange Commission (SEC). 2000. *Final Rule: Audit Committee Disclosure.* Release No. 34-42266; File No. S7-22-99. Available at: http://www.sec.gov/rules/final/34-42266.htm.

Skinner, D. J. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32 (1): 38–60.

Whitehead, J. C., I. M. Millstein, J. H. Biggs, F. J. Borelli, C. A. Bowsher, D. D. Dammerman, R. A. Grasso, P. Laskawy, J. J. Schiro, W. C. Steere, Jr., and F. G. Zarb. 1999. *Report and Recommendations of the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees.* New York, NY: NYSE and NASD.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 779–799

Exploring the Term of the Auditor-Client Relationship and the Quality of Earnings: A Case for Mandatory Auditor Rotation?

James N. Myers
Linda A. Myers
University of Illinois at Urbana-Champaign

Thomas C. Omer
University of Illinois at Chicago

ABSTRACT: In this study, we document evidence on the relation between auditor tenure and earnings quality using the dispersion and sign of both absolute Jones-model abnormal accruals and absolute current accruals as proxies for earnings quality. Our study is motivated by calls for "mandatory auditor rotation," which are based on concerns that longer auditor tenure reduces earnings quality. Multivariate results, controlling for firm age, size, industry growth, cash flows, auditor type (Big N versus non-Big N), industry, and year, generally suggest *higher* earnings quality with longer auditor tenure. We interpret our results as suggesting that, in the current environment, longer auditor tenure, on average, results in auditors placing greater constraints on extreme management decisions in the reporting of financial performance.

Keywords: *auditor tenure; earnings quality; audit quality; mandatory rotation.*

Data Availability: *All data used in this study is publicly available.*

I. INTRODUCTION

The issue of earnings quality and the quality of financial statements in general has been the focus of recent Congressional, regulatory, and financial statement user discussions. Recent events have focused the public's eye on earnings quality and on audit quality in general. Numerous legislative proposals at federal and state levels would place limits on the maximum length of a given auditor-client relationship. The proposed limits are based on the notion that extended auditor tenure results in auditor complacency about and possibly complicity in the decisions that management makes regarding the presentation of financial results. While the Sarbanes-Oxley Act of 2002 does not impose mandatory rotation on audit firms, it mandates a study of mandatory rotation by the Comptroller General of the United States to be completed within one year of the passage of the Act.

We thank Mark Bradshaw, Mark Peecher, Marjorie Shelley, and two anonymous reviewers for their insightful comments. All errors and omissions are our own.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted June 2002
Accepted February 2003

In this study, we document the relation between earnings quality (using absolute, signed, and raw [unsigned] accruals measures as proxies for earnings quality) and auditor tenure. We assert that earnings quality can be used to draw inferences about audit quality. We do not purport to comment on audit quality in the traditional sense—compliance with Generally Accepted Auditing Standards (GAAS). Rather, we claim that when audit quality is high, auditors constrain the extreme (and presumably self-serving) choices that management would like to make in presenting the financial position of the firm. When audit quality is low, we posit that auditors do not constrain these extreme choices and that in some cases, auditors may even aid management in "pushing the boundaries" of Generally Accepted Accounting Principles (GAAP).[1]

Proponents of mandatory rotation express the belief that poor-quality earnings are associated with extended auditor tenure. Poor-quality earnings are problematic because they can mislead investors, resulting in misallocated resources. In its defense, the accounting profession has argued that mandatory rotation increases audit start-up costs and increases the risk of audit failure because the incoming auditor places increased reliance on the client's estimates and representations in the initial years of an engagement. Over time, auditors can gain firm-specific expertise, which helps them to understand the business and allows them to rely less on management estimates.[2] Managers of firms being audited are also opposed to mandatory rotation because changing auditors is costly. These managers believe that auditors better understand their particular business with experience and managers have concerns about whether a new auditor will have the requisite industry expertise and/or will be able to put forth the additional effort required to audit a new client (Dunham 2002).

In this paper, we provide evidence on the claim made by proponents of mandatory auditor rotation that the current voluntary rotation system is flawed. Because the push for mandatory rotation stems from the belief that extended auditor tenure impairs auditor independence, we assert that understanding the relation between auditor tenure and earnings quality is central to the debate. That is, the veracity of claims made by proponents about the current system should be central to decisions made by regulators, and of interest to auditors and financial statement users. To our knowledge, this is the first broad cross-sectional study to explicitly consider the relation between auditor tenure and earnings quality as proxied for by measures of accounting accruals. We assume that producing higher-quality earnings is the goal of recent legislative proposals and claim that the extent to which auditor tenure either enhances or inhibits earnings quality should be addressed before legislative or regulatory changes impose limits on auditor tenure. Borrowing extensively from previous research that uses accounting accruals as a proxy for earnings quality, we test for an association between auditor tenure and earnings quality for those auditor-client combinations where the auditor-client relationship lasts for at least five years. We do this by examining the relation between absolute, signed, and raw values of both Discretionary and Current Accruals and auditor tenure after controlling for other factors expected to affect the distribution of accruals.

Although we cannot comment on whether mandatory auditor rotation would improve earnings quality, we do provide evidence inconsistent with the claim that earnings quality deteriorates with extended auditor tenure under the current voluntary rotation system. After controlling for firm age, size, industry growth, cash flows, auditor type (Big N versus non-Big N), industry, and year, we find that the magnitude of both Discretionary and Current Accruals decline with longer auditor tenure. Further, we find some evidence that longer auditor tenure is associated with *less* extreme income-increasing accruals.[3] This suggests that as the relationship lengthens, auditors limit management's ability to use accruals to increase current period earnings. We also find evidence that longer auditor tenure is associated with *less* extreme income-decreasing accruals. This suggests that as the relationship lengthens, auditors limit management's ability to create reserves to manage future earnings. We find this result particularly interesting because income-decreasing earnings management has received a great deal of attention from regulators and the popular press as of late.[4]

---

[1] For a more thorough discussion of audit quality and auditor independence, see Levitt (2000).

[2] For example, while Chairman of the SEC, Wallman wrote that "[p]eriodically rotating the audit firms of a public company seems contrary to the notion of learning as much as possible about an audit client. It also would appear to be remarkably inefficient." (See Wallman, 1996, 93.)

The remainder of the paper is organized as follows. In Section II, we summarize the arguments supporting and opposing mandatory auditor rotation and develop our hypotheses. We discuss our data and empirical results in Section III, and in Section IV we present our conclusions and the implications of our study.

II. HYPOTHESIS DEVELOPMENT

Arguments Supporting Mandatory Auditor Rotation

Whether auditor rotation should be made mandatory is an issue that has been debated for more than 40 years in the U.S. Proponents of mandatory auditor rotation are generally concerned that auditor independence, and thus audit quality, will decrease with increased auditor tenure. Mautz and Sharaf (1961) suggest that extended auditor-client relationships could have a detrimental affect on auditor independence because an auditor's objectivity about a client is reduced with the passage of time. Further, the Metcalf Committee states that "mandatory auditor rotation is a way to bolster auditor independence" (U.S. Senate 1976, 21). More recently, regulators suggest a link between auditor tenure and reductions in earnings quality, and allude to mandatory auditor rotation as a possible solution (U.S. Senate 1977; AICPA 1978; Berton 1991; SEC 1994). It has been suggested that decreased auditor independence could lead to auditors' support for more aggressive accounting choices that "push the boundaries" of GAAP and could ultimately result in a failure to detect material fraud and/or misstatements. Mandatory rotation proponents argue that limiting auditor tenure reduces concerns about deteriorating independence and audit quality.[5] In response to the recent Enron/Anderson audit failure, several bills containing provisions limiting auditor tenure and mandating auditor rotation were proposed in the House and Senate as part of an effort to improve financial reporting and protect investors.[6]

---

[3] Specifically, we find a significantly negative association between tenure and Positive Current Accruals in all specifications. We also find a significantly negative association between tenure and Positive Discretionary Accruals when all sample firm-years are included in our analyses, but when we omit firms with extremely positive and negative return on assets, the sign on tenure remains negative but the result is not significant.

[4] See, for example, Levitt (1998).

[5] Research by Deis and Giroux (1992) suggests a negative relation between auditor tenure and audit quality, consistent with concerns about the current system raised by proponents of mandatory auditor rotation. However, their results may have little bearing on the current debate because their sample includes only small CPA firms and the audit clients represent quasi-governmental entities in the public sector.

[6] See, for example, the Integrity in Auditing Act of 2002 (which suggests that auditors should not be considered independent if they have audited a firm for more than seven consecutive years), the Comprehensive Investor Protection Act of 2002 (which suggests that non-independence is a problem when the auditor has consecutively audited a firm's financial statements for more than four years), and the Truth and Accountability in Accounting Act of 2002 (which states that an auditor should not be considered independent if it has audited a firm's financial statements for more than seven consecutive years).

Arguments Opposing Mandatory Rotation

Opponents of mandatory rotation suggest that along with higher audit costs, an increased likelihood of audit failures would result from mandated rotation. New auditors, they argue, lack sufficient knowledge regarding firm-specific risks and, as a consequence, audit failures would likely increase. This argument is consistent with research indicating that a greater proportion of audit failures occur on newly acquired clients (Berton 1991; Petty and Cuganesan 1996) and that auditors' litigation risk is greater in the early years of an engagement (Palmrose 1986b, 1991). Further evidence on auditor tenure and audit quality is provided by the AICPA's Quality Control Inquiry Committee of the SEC Practice Section. The Committee analyzed 406 cases of alleged audit failure between 1979 and 1991 and concluded that *allegations* of audit failure occur almost three times as often when an audit firm is performing its first or second audit of a given client (AICPA 1992). Finally, the accounting profession argues that uncertainty regarding characteristics of the client increases the potential for audit failures early in the auditor-client relationship (PricewaterhouseCoopers 2002).[7]

Accounting Accruals and Audit Quality

Absent from public discussion and prior research is a cogent picture of the (inter)temporal relation between audit quality and auditor tenure. The results of prior studies, while informative about extreme events (for example, SEC violations and qualified opinions), do not address audit quality for a broad cross-section of firms. However, the relation between auditor tenure and audit quality is of interest in a broader context. By linking earnings quality and audit quality, we attempt to address this relation more generally. While we acknowledge that studying the association between auditor tenure and audit quality in the case of extreme audit failures is important, we suggest that a broader analysis of this relation should be informative to the mandatory rotation debate.

To address the relation between auditor tenure and audit quality, we must first provide a plausible definition of audit quality. We borrow from an extensive prior literature to argue that accounting accruals measures are a plausible descriptor of audit quality. Numerous studies have examined the association between various measures of accruals and proxies for audit quality. These proxies include auditor litigation, qualified audit opinions, audit failures, and auditor conservatism. Results suggest that higher accruals levels are positively associated with eventual auditor litigation (Heninger 2001), the issuance of qualified audit opinions (Bartov et al. 2000), audit failures (Geiger and Raghunandan 2002), and auditor changes (DeFond and Subramanyam 1998), while lower accruals levels are associated with greater auditor conservatism, which has been interpreted as suggestive of higher-quality audits (Becker et al. 1998; Francis et al. 1999; Francis and Krishnan 1999). Based on these findings, we posit that high-quality audits mitigate more extreme management reporting decisions, and suggest that accruals can be used to identify these extreme reporting decisions.

---

[7] Prior research suggests two reasons why auditors are likely to be less aggressive in their oversight of management early in an auditor's tenure. First, Geiger and Raghunandan (2002) suggest that one reason new auditors are less aggressive in their oversight is because they are attempting to recoup losses from the competitive practice of low-balling. However, while the practice of low-balling has been documented by Simon and Francis (1988), Ettredge and Greenberg (1990), and Deis and Giroux (1996), the literature suggests that low-balling and the recouping of losses does not impair auditor independence (DeAngelo 1981, 113; Lee and Gu 1998, 534). Thus, if auditor independence is not impaired, auditor oversight of managements' decisions should not be affected regardless of tenure. Second, Geiger and Raghunandan (2002) also suggest that new auditors are likely to be less aggressive because they lack knowledge of client-specific risks. Although they find that auditors are less likely to issue going concern opinions for new clients, this does not provide convincing evidence on either the relation between low-balling and audit quality or the lack of client-specific knowledge. That is, their result can be interpreted that the auditors failed to issue going concern opinions to increase the likelihood that they would recoup some or all of the fees lost in the initial engagement period *or* that new auditors were unaware of the client-specific risks and this lack of knowledge lead to the issuance of nonqualified opinions.

Accruals measures have also been used in numerous accounting studies that consider the relation between earnings management and accruals behavior, and earnings quality and accruals behavior. Early studies used the change in Total Accruals as a measure of management discretion (e.g., Healy 1985; DeAngelo 1986), while later studies adopted and modified the "Jones Model" (Jones 1991).[8] Thus, abnormal discretionary accruals (estimated using variations of the Jones Model) became the accepted proxy for extreme managerial discretion. Various accruals measures have also been used to study earnings quality. These studies are important to the audit quality issue because they generally find that information is contained in specific accruals and/or that earnings quality declines with extreme accruals (e.g., Sloan 1996; Thomas and Zhang 2001; Xie 2001; Dechow and Dichev 2002; Richardson et al. 2002). Both the literatures on earnings management and earnings quality provide evidence that suggests that extreme accruals are less desirable, consistent with the audit quality literature cited above. Thus, the extent to which accruals behavior is associated with auditor tenure under the current voluntary rotation system is informative to parties currently debating mandatory rotation.

We follow prior literature and perform analyses of those accruals measures generally considered to be most representative of managements' discretion in the financial reporting process—Discretionary Accruals and Current Accruals. As in prior work, we not only study the raw and absolute values of these measures, but we also posit that studying the association between auditor tenure and *signed* accruals is important because auditors could constrain income-increasing accruals to a greater or lesser extent than income-decreasing accruals, and auditor tenure could have differential effects on this constraint.[9] By using multiple measures, we better test the relation between auditor tenure and audit quality.

We formulate the following alternative hypotheses regarding the relation between auditor tenure and accounting accruals:

$H_{\text{rotation proponent}}$: Audit quality (as measured by accounting accruals) is decreasing in auditor tenure.

$H_{\text{rotation opponent}}$: Audit quality (as measured by accounting accruals) is increasing in auditor tenure.

We test these general hypotheses to determine which is descriptive of the relation between auditor tenure and audit quality (as we have defined it).[10]

---

[8] For an extensive review of the earnings management literature, see Healy and Wahlen (1999) and Dechow and Skinner (2000).

[9] For example, the perception that litigation is more likely to occur when income is overstated might lead auditors to constrain income-increasing accruals more severely than income-decreasing accruals even if auditors become less independent of their clients as the relationship lengthens. Tests on the signed accruals rather than on their absolute values allow us to examine this possibility.

[10] In related work, Davis et al. (2002) provides results suggesting that discretionary accruals increase with auditor tenure. However, differences in design (noted below) make it difficult to compare this study with ours. First, we examine the absolute value of two accruals measures and separately examine the positive and negative components of these measures. Second, we do not require that firms survive throughout our entire sample period, thus eliminating the survivorship bias in Davis et al. (2002). Third, we delete observations that limit our ability to generalize our findings. These include: (1) very young firms, so that abnormal accruals behavior associated

(continued on next page)

III. DATA AND EMPIRICAL RESULTS

Data

Our initial sample consists of all firm-years from 1988 to 2000 inclusive with sufficient data on the 2001 Compustat annual industrial and research files to estimate accruals. In addition, we require at least the six years of prior data to ensure that any abnormal accruals behavior associated with start-up firms (Teoh et al. 1998a, 1998b) is not attributed to short auditor tenure. Because Collins and Hribar (2002) show that estimating accruals is problematic when firms undergo mergers and acquisitions, we also omit these observations.[11] Finally, we limit our analyses to years beginning in 1988 because cash flow from operations was not available before that date.[12]

We calculate auditor tenure as the number of years that the firm has retained the given auditor, and code auditor changes attributable to audit firm mergers as a continuation of the prior auditor. We determine which observations are in the top or bottom 0.5 percent of the distribution of cash flow from operations, Current Accruals or Jones Model residuals annually and delete them to remove outliers. Note that our research question relates to the relation between the accruals (for a given firm) and auditor tenure. To guard against the possibility that observations where firms switch auditors early in the relationship differ systematically from observations where firms retain the auditor for a number of years, we eliminate all firm-year observations where the auditor-client relationship did not last at least five years. That is, when the auditor-client relationship lasts at least five years, we retain all available observations, and when the auditor-client relationship lasts for less than five years, we eliminate all of these auditor-client observations. In this way, the early and later years represent the same auditor-client combinations, thus preserving comparability. Eliminating the quick turnover group reduces the sample size to 42,302 firm-years.

In our analyses, we use two measures of accruals:

$$\mathit{Current\ Accruals} = ((\Delta CA - \Delta Cash) - (\Delta CL - \Delta STD))$$

and

$$\mathit{Discretionary\ Accruals} = \mathit{Accruals} - \beta_0 + \beta_1\,\Delta Revenue + \beta_2 PPE$$
(Jones_Model_residuals)[13]

where:

ΔCA = change in current assets (Compustat annual data item 4);
$\Delta Cash$ = change in cash and cash equivalents (Compustat annual date item 1);

Footnote 10, continued
with start-up firms (Teoh et al. 1998a, 1998b) is not attributed to short auditor tenure; (2) firms undergoing mergers and acquisitions, so that the difficulties in reliably estimating accruals for these firms (Collins and Hribar 2002) do not influence our findings; and (3) observations where firms change auditors in the first five years of an audit engagement, because their inclusion does not allow us to directly address the question of what happens to accruals with lengthy auditor tenure.

[11] To identify firms that undergo mergers and acquisitions, we follow Collins and Hribar (2002) and use the footnote codes appearing in Compustat.

[12] The number of new auditor-client combinations each year ranges between 306 (in 1994) and 527 (in 2000). The number of new auditor-client combinations for Big N auditors is 3,590 versus 1,561 for non-Big N auditors.

[13] We estimate the residuals and the coefficients contemporaneously. That is, for each year, we estimate the model within two-digit SIC industries, and use the residuals from this regression as our estimate of discretionary accruals.

ΔCL = change in current liabilities (Compustat annual data item 5);
ΔSTD = change in short-term notes and current portion of long-term debt (Compustat annual data item 34);
β_i = coefficient estimates from an OLS regression by industry-year, where an industry is defined by its two-digit SIC code; to reduce sampling error and to be consistent with prior literature, we eliminate industry-years with fewer than eight firms;
$Accruals$ = operating income after depreciation (Compustat annual data item 178) – cash flow from operations (Compustat annual data item 108);
$\Delta Revenue$ = change in net sales revenue (Compustat annual data item 12); and
PPE = property, plant, and equipment—net (Compustat annual data item 8).

We scale all variables by average total assets.

Empirical Results

Table 1, Panel A provides descriptive statistics for all firms in our sample where the auditor-client relationship lasts at least five years, while Panel B provides descriptive statistics for the quick-turnover firms (where the auditor-client relationship lasts for less than five years). Within each panel we present descriptive statistics during three stages of the auditor-client relationship: 1 to 2 years; 3 to 4 years; and more than 4 years (when an auditor-client relationship lasts for at least 5 years).[14] We selected these stages (or groupings) for two reasons. First, drafts of the Congressional bills cited earlier provided a range of suggested mandatory rotation periods, and groups of 3 to 5 years and more than five years are representative of the suggested periods. Second, the argument for mandatory auditor rotation hinges, in part, on the potential for reducing extreme management reporting decisions with unbiased oversight by auditors in the initial years of an auditor-client relationship, while the argument against mandatory auditor rotation hinges, in part, on the uncertainty that occurs during the initial years. Our grouping of 1 to 2 years is representative of this initial period.

The quick-turnover firms are significantly different from firms where the auditor-client relationship lasts for at least five years (i.e., the sample firms). For example, for the subset of firms that are in their first or second year of the audit engagement, the quick-turnover firms have significantly smaller average ROA and Cash Flow, and are older. Results on size proxies are mixed. Table 1 also reveals that the dispersion of our two accruals measures is significantly greater (in the first two years of the audit engagement) for the quick-turnover firms than for the sample firms. For these reasons, and because including quick-turnover firms does not allow us to adequately address the possible consequences of extended auditor tenure, we eliminate the quick-turnover firms in all subsequent analyses.

Univariate Results

In this section, we present univariate results concerning auditor tenure and our accruals measures. Specifically, in Figure 1 we graph the 90th percentile, mean, median, and 10th percentile of our two accruals measures, Discretionary Accruals and Current Accruals, by

---

[14] Note that observations relating to a given auditor-client combination will appear in Panel A if and only if that combination lasted for at least five years. However, if the auditor-client combination lasted for ten years, the first two years will be included in the Group 1 observations, the next two years will be included in the Group 2 observations, and the final six years will be included in the Group 3 observations. Observations relating to a given auditor-client combination will appear in Panel B if and only if that combination lasted for less that five years. The groupings can be interpreted as in Panel A.

TABLE 1
Descriptive Statistics by Tenure Group

Tenure Group	ROA Mean (Median) [n]	Cash Flow Mean (Median) [n]	MVE Mean (Median) [n]	Assets Mean (Median) [n]	Tenure Mean (Median) [n]	Age Mean (Median) [n]	Absolute Discretionary Accruals Mean (Median) [n]	Absolute Current Accruals Mean (Median) [n]
Panel A: Omitting Observations Where the Auditor-Client Relationship Lasts for Less than Five Years								
Group 1	−0.0324	0.0394	781.11	1,153.58	1.6011	10.7144	0.0918	0.1582
1 or 2	(0.0237)	(0.0635)	(32.868)	(41.201)	(2)	(10)	(0.0553)	(0.0702)
years	[2,637]	[2,637]	[2,174]	[2,637]	[2,637]	[2,637]	[2,637]	[2,637]
Group 2	−0.0407	0.0293	991.15	1,463.79	3.5357	11.4434	0.0917	0.1565
3 or 4	(0.0249)	(0.0631)	(41.135)	(49.897)	(4)	(11)	(0.0524)	(0.0677)
years	[3,730]	[3,730]	[3,115]	[3,730]	[3,730]	[3,730]	[3,730]	[3,730]
Group 3	−0.0193	0.0477	1,788.15	2,329.66	10.4634	12.1084	0.0812	0.1313
>4 years	(0.0316)	(0.0723)	(93.283)	(129.472)	(10)	(11)	(0.0484)	(0.0592)
	[35,935]	[35,935]	[31,124]	[35,935]	[35,935]	[35,935]	[35,935]	[35,935]
Panel B: Observations Where the Auditor-Client Relationship Lasts for Less than Five Years								
Group 1	−0.0939	0.0017	1,024.55	1,721.13	1.3494	11.4566	0.1091	0.1763
1 or 2	(0.0077)	(0.0457)	(29.176)	(46.152)	(1)	(10)	(0.0622)	(0.0793)
years	[6,844]	[6,844]	[5,851]	[6,844]	[6,844]	[6,844]	[6,844]	[6,844]
Group 2	−0.1048	−0.0080	827.15	1,246.54	3.3087	11.6249	0.1163	0.1992
3 or 4	(0.0043)	(0.0388)	(22.586)	(33.953)	(3)	(11)	(0.0723)	(0.0817)
years	[1,869]	[1,869]	[1,594]	[1,869]	[1,869]	[1,869]	[1,869]	[1,869]

(continued on next page)

TABLE 1 (continued)

Tenure Group	ROA Mean (Median) [n]	Cash Flow Mean (Median) [n]	MVE Mean (Median) [n]	Assets Mean (Median) [n]	Tenure Mean (Median) [n]	Age Mean (Median) [n]	Absolute Discretionary Accruals Mean (Median) [n]	Absolute Current Accruals Mean (Median) [n]
Panel C: Tests of Differences between Observations Where the Auditor-Client Relationship Lasts at Least Five Years and Observations Where the Auditor-Client Relationship Lasts for Less than Five Years								
t-statistics for Tests of Differences in Means (Z-statistics for Tests of Differences in Medians)								
Group 1 1 or 2 years	7.04*** (8.65***)	7.91*** (5.90***)	−1.98* (1.76)	−3.13*** (−1.71)	22.59*** (22.25***)	−8.84*** (−1.37)	−6.55*** (−3.27)***	−2.26** (−3.64)***
Group 2 3 or 4 years	5.90*** (8.35***)	5.13*** (6.82***)	0.97 (6.64***)	0.83 (4.78***)	16.87*** (16.07***)	−1.53 (−1.83***)	−6.76*** (−7.52***)	−3.62*** (−3.78***)

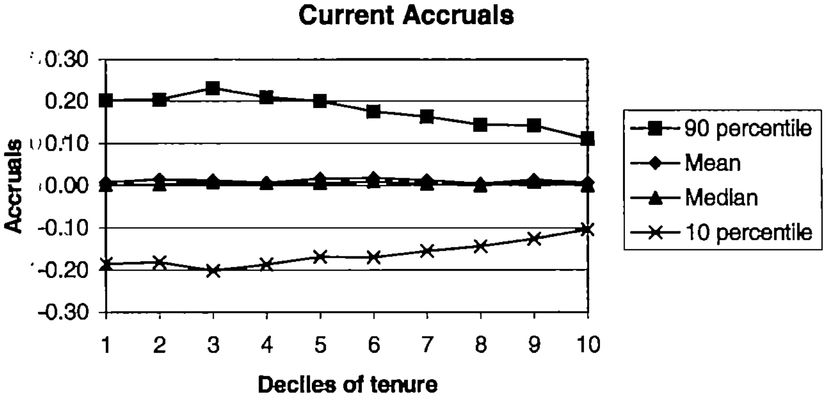
---

*, **, *** Indicates significant at .10, .05, and .01, respectively, using a two-tailed test.

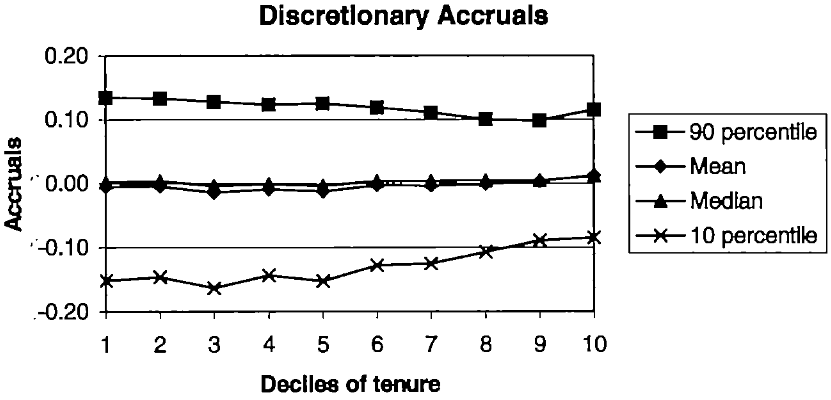
Observations relating to a given auditor-client combination will appear in Panel A if and only if that combination lasted for at least five years. However, if the auditor-client combination lasted ten years, the first two years will be included in the Group 1 observations, the next two years will be included in Group 2 observations, and the final six years will be included in Group 3 observations. Observations relating to a given auditor-client combination will appear in Panel B if and only if that combination lasted for less that five years. The groupings can be interpreted as in Panel A.

FIGURE 1
Accruals by Deciles of Tenure

Panel A: Current Accruals by Deciles of Tenure



Panel B: Discretionary Accruals by Deciles of Tenure



Omitting observations where the auditor-client relationship lasts for less than five years.

$$Current\ Accruals = ((\Delta CA - \Delta Cash) - (\Delta CL - \Delta STD))$$

where:

ΔCA = change in current assets (Compustat annual data item 4);
$\Delta Cash$ = change in cash and cash equivalents (Compustat annual date item 1);
ΔCL = change in current liabilities (Compustat annual data item 5); and
ΔSTD = change in short-term notes and current portion of long-term debt (Compustat annual data item 34).

$$Discretionary\ Accruals = Accruals - \beta_0 + \beta_1 \Delta Revenue + \beta_2 PPE$$

where:

$Accruals$ = operating income after depreciation (Compustat annual data item 178) – cash flow from operations (Compustat annual data item 108);
β_i = coefficient estimates from an OLS regression by industry-year, where an industry is defined by its two-digit SIC code;
$\Delta Revenue$ = change in net sales revenue (Compustat annual data item 12); and
PPE = property, plant, and equipment—net (Compustat annual data item 8).

All variables are scaled by average total assets.

tenure decile. Here, the 90th percentile represents extreme positive accruals while the 10th percentile represents extreme negative accruals.[15]

Results in Panel A of Figure 1 suggest that tenure has little effect on the average reported Current Accruals. That is, mean and median Current Accruals do not appear to differ in the early years and later years of tenure. However, tenure does seem to have significant affect on the extremes of the distribution. The extreme deciles of Current Accruals decline fairly steadily with tenure. In fact, after the third decile of tenure, the effect is monotonic.

Results in Panel B of Figure 1, which examine the relation between Discretionary Accruals and auditor tenure, are similar to those in Panel A, but the affect of tenure on large positive Discretionary Accruals is more ambiguous. That is, the monotonic results for Current Accruals are not evident for large positive Discretionary Accruals. Further, the mean and median Discretionary Accruals are slightly higher for firms with long tenure than for firms with short tenure. The most striking finding, however, is the sharp decline in extreme negative Discretionary Accruals as tenure increases.

Overall, Figure 1 reveals that the negative values of both accruals measures are less extreme (i.e., closer to zero) when auditor tenure is greater. Further, positive Current Accruals also exhibit a similar characteristic. This suggests that as tenure lengthens, auditors constrain management's ability to make extreme reporting decisions.[16]

While these results are consistent with the notion that managers record less extreme accruals as auditor tenure lengthens, we have not yet incorporated other factors that have been shown to influence the cross-sectional distribution of accruals. To ensure that auditor tenure is not a proxy for one or more of these factors, we perform multivariate analyses in the next section.

Multivariate Results

In this section, we document the general association between auditor tenure and accruals. Because our interest is in the cross-section, rather than in cases where extreme audit failures occur, we do not condition on events suggested by prior literature as being associated with extreme accruals. With that in mind, we suggest the following general model.[17]

$$\begin{aligned} Accruals = {} & \alpha + \beta_1 Tenure + \beta_2 Age + \beta_3 Size + \beta_4 IndustryGrowth \\ & + \beta_5 CashFlow + \beta_6 AuditorType + \beta_{7\text{–}69} Industry \\ & + \beta_{1988\text{–}1999} Year + \varepsilon \end{aligned} \tag{1}$$

where:

Accruals = Discretionary Accruals or Current Accruals (measured in raw, absolute, and positive and negative values) scaled by average total assets;

Tenure = the number of consecutive years that the firm has retained the auditor;

---

[15] We group firms by tenure decile (rather than by year of tenure) because there are fewer observations with very long tenure (i.e., greater than 15 years) than with very short tenure (i.e., 1 to 5 years) leading to differing precision levels for the accruals estimates of short- and long-tenure firms. A small number of very long-tenure firms would lead to high variance in the different measures of accruals. Aggregating into deciles mitigates this problem. The lower bound for each decile (1 through 10) is 1, 3, 5, 6, 7, 9, 10, 11, 13, and 16 years, respectively.

[16] Further, the mean absolute Discretionary (Current) Accruals as a percentage of Total Assets decline with auditor tenure. For the 1st, 5th, and 10th deciles of auditor tenure, the mean absolute Discretionary (Current) Accruals are 0.09 (0.16), 0.09 (0.14), and 0.07 (0.08), respectively.

[17] Firm and year subscripts are suppressed for presentation.

Age = the number of years for which total assets was reported in Compustat since 1980;
Size = the log of total assets;
Industry Growth = $\sum_{i=1}^{N} Sales_{i,t} / \sum_{i=1}^{N} Sales_{i,t-1}$ by two-digit SIC code;
Cash Flow = the firm's cash flow from operations divided by average total assets;
Auditor Type = an indicator set to 1 if the auditor is a Big N Auditor, and set to 0 otherwise;[18]
Industry = a control based on two-digit SIC codes (62 industries); and
Year = a control for calendar year.

We are primarily concerned with the sign of the coefficient on *Tenure*. If proponents of mandatory rotation are correct, then we expect the coefficient on *Tenure* to be positive (suggesting that auditor independence becomes impaired with longer auditor tenure). However, if opponents of mandatory rotation are correct, then we expect the coefficient on *Tenure* to be negative (suggesting that auditors place greater constraints on management reporting decisions with greater auditor tenure).

We include control variables in the model to reduce the possibility that our auditor tenure results proxy for other cross-sectional determinants of accruals. Specifically, we include *Age* because accruals differ with changes in firm life cycle (Anthony and Ramesh 1992; Dechow et al. 2001). *Size* is included because large firms tend to record larger, more stable accruals (Dechow and Dichev 2002). We control for *Industry Growth* (measured as growth in industry sales) because growth in the industry should be positively correlated with the accruals. *Cash Flow* is included because firms with higher cash flow from operations are more likely to be better performers (Frankel et al. 2002) and because and accruals and cash flows are negatively correlated on average (Dechow 1994; Sloan 1996). Because prior research suggests that large audit firms tend to be more conservative and their conservatism tends to limit extreme accruals (Becker et al. 1998; Francis et al. 1999; Francis and Krishnan 1999), we include *Auditor Type*. Finally, we include *Industry* because the types of accruals vary by industry (Barth et al. 2001) and the magnitude and type could vary by year.

When performing our analyses, we acknowledge that auditors might require managers to record *legitimate* asset write-offs and/or impairment charges in the initial year of an auditor-client relationship. These write-offs and charges would create large absolute accruals and large negative accruals in the first year, but we posit that these charges are not earnings management. To ensure that these charges do not influence our interpretation of the relation between auditor tenure and audit quality, we eliminate the first year of the audit in all of our multivariate tests. We estimate Model (1) in Table 2 using all remaining sample observations. We present results on the absolute values in Panel A; on the signed (positive and negative) values in Panel B, and on the raw values in Panel C for our two accruals measures. Panel A reveals that the coefficients for *Tenure* are negative and significant for both Discretionary and Current Accruals. That is, extended auditor tenure reduces the dispersion of both measures, consistent with the notion that longer auditor tenure results in higher audit quality.

However, regulators are not concerned solely with the dispersion in accruals. Rather, they are concerned about the distortion in earnings that can occur because of inappropriate

---

[18] This measure varies with mergers of the large accounting firms; the Big N is the largest number of major accounting firms considered.

TABLE 2
Regressions of Accruals Measures on Auditor Tenure and Control Variables
All Observations in the Sample

$$Accruals = \alpha - \beta_1 Tenure + \beta_2 Age + \beta_3 Size + \beta_4 IndustryGrowth + \beta_5 CashFlow + \beta_6 AuditorType + \beta_{7-69} Industry + \beta_{1988-1999} Year + \varepsilon$$

Panel A: Absolute Value Measures

Variable	Discretionary Accruals	Current Accruals
Intercept	0.19275	0.25324
	[< 0.0001]	[< 0.0001]
Tenure	−0.00037	−0.00142
	[0.0041]	[0.0004]
Age	−0.00038	−0.00393
	[0.0121]	[< 0.0001]
Size	−0.01148	−0.01032
	[< 0.0001]	[< 0.0001]
Industry Growth	−0.00305	0.00432
	[0.2348]	[0.5807]
Cash Flow	−0.05453	−0.18027
	[< 0.0001]	[< 0.0001]
Auditor Type	−0.00356	0.01283
	[0.0146]	[0.0039]
n	41,250	41,250

Panel B: Signed Measures

Variable	Positive Discretionary Accruals	Negative Discretionary Accruals	Positive Current Accruals	Negative Current Accruals
Intercept	0.19450	−0.5468	−0.2555	0.6487
	[< 0.0001]	[0.4689]	[< 0.0001]	[0.7444]
Tenure	−0.00146	0.0534	−0.0293	0.2463
	[0.0502]	[< 0.0001]	[< 0.0001]	[< 0.0001]
Age	0.00004	0.0803	−0.0526	0.4188
	[0.9610]	[< 0.0001]	[< 0.0001]	[< 0.0001]
Size	−0.09200	0.8027	−0.1102	1.1682
	[< 0.0001]	[< 0.0001]	[< 0.0001]	[< 0.0001]
Industry Growth	−0.00830	1.0595	0.0986	0.1010
	[< 0.0001]	[0.0006]	[0.0284]	[0.7942]
Cash Flow	−0.11950	−0.1302	−0.6248	0.9005
	[< 0.0001]	[0.0709]	[< 0.0001]	[< 0.0001]
Auditor Type	−0.00882	−0.4710	0.0658	−0.7464
	[0.2650]	[< 0.0001]	[0.0124]	[< 0.0001]
n	21,449	19,801	21,619	19,631

(continued on next page)

TABLE 2 (continued)

Panel C: Raw Measures

Variable	Discretionary Accruals	Current Accruals
Intercept	−0.03712 [0.0039]	−0.00154 [0.9619]
Tenure	0.00038 [0.0330]	−0.00043 [0.3359]
Age	0.00269 [< 0.0001]	0.00009 [0.8581]
Size	0.00077 [0.0228]	0.00387 [< 0.0001]
Industry Growth	0.00142 [0.6869]	−0.00121 [0.8914]
Cash Flow	−0.07104 [< 0.0001]	−0.10073 [< 0.0001]
Auditor Type	−0.02205 [< 0.0001]	−0.01294 [0.0103]
n	41,250	41,250

Each cell contains a coefficient estimate and p-values in square brackets.
Omitting observations where the auditor-client relationship lasts for less than five years, observations where firms reported a merger or acquisition in the year, and the first year of an auditor-client relationship.

Accruals = one of the accruals measures presented in Figure 1;
Tenure = the number of consecutive years that the firm has retained the auditor;
Age = the number of years of data available on Compustat since 1980;
Size = the log of total assets;
Industry Growth = $\sum_{i=1}^{N} Sales_{i,t} / \sum_{i=1}^{N} Sales_{i,t-1}$ by two-digit SIC code;
Cash Flow = the firm's cash flow from operations divided by average total assets;
Auditor Type = an indicator set to 1 if the auditor is a Big N auditor, and set to 0 otherwise;
Industry = a control based on two-digit SIC codes (62 industries); and
Year = a control based on calendar year.

income-increasing accruals and/or inappropriate income-decreasing accruals. Income-increasing accruals can be used to inflate current earnings, while income-decreasing accruals can be used to create "cookie jar reserves," which allow managers to increase future earnings. Whether extended auditor tenure constrains either or both of these types of accruals should be informative to regulators. To address this, we estimate Model (1) using the signed values for each accruals measure in Panel B of Table 2.[19] The results indicate that extended auditor tenure constrains both income-increasing and income-decreasing accruals. That is, the coefficient for *Tenure* is negative and significant for positive accruals and positive and significant for negative accruals. This suggests that auditor oversight of management decisions is generally balanced in that auditors appear to constrain extreme income-increasing and income-decreasing accruals.

In Table 2, Panel C we estimate Model (1) using the raw values of our accruals measures. The results are mixed and highlight the need to conduct the analyses presented in

---

[19] Here we use a truncated regression approach because the ordinary least squares (OLS) estimates are generally biased toward zero when a sample is truncated. Thus, we estimate a maximum likelihood truncated regression model to arrive at an unbiased coefficient on *Tenure*. See Greene (2000, 682–690) or Maddala (1977, 269–273) for a description of the truncated regression approach.

Panels A and B. The coefficient for *Tenure* using Discretionary Accruals is significant and positive suggesting that as auditor tenure increases, the level of Discretionary Accruals increases. This is consistent with arguments made by proponents of mandatory rotation. The coefficient for *Tenure* using Current Accruals is negative but not significant, suggesting no relation between auditor tenure and the level of Current Accruals. However, we posit that regulators (and users of financial statements) should not be as concerned about the level of accruals as about the extremes of the distribution and the directional affect. Thus, looking only at levels (in Panel C) provides a less-than-satisfactory indication of the relation between auditor tenure and audit quality.

In sum, the results in Panels A and B of Table 2 suggest that extended auditor tenure has a beneficial effect on the dispersion of accruals (i.e., accruals are less extreme with extended auditor tenure) and that auditors constrain both income-increasing and income-decreasing accruals as auditor tenure increases.

With respect to the control variables, our analyses reveal no consistent results for *Age* or *Industry Growth*, but we do find that as firm size increases, accruals are more positive (consistent with Dechow and Dichev [2002]), and that as firm size increases, both positive and negative accruals are smaller. The coefficient on *Cash Flow* is significantly negative in all specifications (other than for negative Current Accruals, in which case it is positive), consistent with findings in Dechow (1994) and Sloan (1996). Further, the coefficient for *Auditor Type* is negative for all Discretionary Accruals measures but is mixed for Current Accruals measures. Our results for the absolute Discretionary Accruals and raw Discretionary and Current Accruals measures are consistent with prior research suggesting that Big N auditor clients report more conservative earnings than do non-Big N auditor clients. On the other hand, results for the absolute Current Accruals and positive and negative Discretionary and Current Accruals measures suggest that Big N auditors constrain the reporting choices of their clients less than non-Big N auditors do.

Next, we re-estimate Model (1) omitting observations in the extreme (top and bottom) deciles of ROA (measured by year and industry) to control for the possibility that the association between auditor tenure and accruals is the result of poor financial performance early in the auditor's tenure (DeFond and Park 1997; DeFond and Subramanyam 1998). These results are presented in Table 3, Panel A (absolute values), Panel B (signed values), and Panel C (raw values) for our two accruals measures. While our results on *Tenure* for income-decreasing accruals are generally consistent with the results in Table 2, we find insignificant results on income-increasing Discretionary Accruals for the non-extreme ROA subsample. Thus, we conclude that our main results on income-decreasing accruals are not driven by extreme performance, but our main results on income-increasing accruals are influenced by firms with extreme financial performance.

Sensitivity Analyses

To test whether the association between auditor tenure and accruals is affected by our design choices and/or sample composition, we perform a number of additional tests. First, to test whether industry specialization is important, we form an indicator for specialization and interact it with the tenure variable. In this way, we attempt to address whether an auditor's reduced knowledge regarding the client's industry in the early years of an auditor-client relationship[20] has an affect on the association between tenure and earnings quality. Our specialist definition is based on that in Palmrose (1986a). Here, we code an auditor as a (two-digit SIC code) industry specialist if the auditor audits the largest percentage of

---

[20] Recall that this is one of the arguments against mandatory rotation.

TABLE 3
Regressions of Accruals Measures on Auditor Tenure and Control Variables Eliminating Observations with Extreme ROA

$$Accruals = \alpha + \beta_1 Tenure + \beta_2 Age + \beta_3 Size + \beta_4 IndustryGrowth + \beta_5 CashFlow + \beta_6 AuditorType + \beta_{7-69} Industry + \beta_{1988-1999} Year + \varepsilon$$

Panel A: Absolute Value Measures

Variable	Discretionary Accruals	Current Accruals
Intercept	0.10165 [< 0.0001]	0.19663 [< 0.0001]
Tenure	−0.00029 [0.0004]	−0.00093 [0.0034]
Age	−0.00066 [< 0.0001]	−0.00296 [< 0.0001]
Size	−0.00608 [< 0.0001]	−0.00542 [< 0.0001]
Industry Growth	−0.00093 [0.5687]	0.00580 [0.3524]
Cash Flow	−0.03480 [< 0.0001]	−0.31820 [< 0.0001]
Auditor Type	0.00170 [0.0764]	0.00570 [0.1208]
n	37,571	37,571

Panel B: Signed Measures

Variable	Positive Discretionary Accruals	Negative Discretionary Accruals	Positive Current Accruals	Negative Current Accruals
Intercept	−0.01210 [0.6106]	−0.03000 [0.7367]	−0.2749 [< 0.0001]	8.9784 [0.0106]
Tenure	−0.00012 [0.7020]	0.00364 [0.0021]	−0.0367 [< 0.0001]	0.1771 [< 0.0001]
Age	0.00153 [< 0.0001]	0.01160 [< 0.0001]	−0.0536 [< 0.0001]	0.4614 [< 0.0001]
Size	−0.01200 [< 0.0001]	0.06760 [< 0.0001]	−0.0868 [< 0.0001]	0.8708 [< 0.0001]
Industry Growth	0.01450 [0.0590]	0.08300 [0.0044]	0.1406 [< 0.0001]	0.5380 [0.2992]
Cash Flow	−0.61090 [< 0.0001]	−0.82370 [< 0.0001]	−2.4576 [< 0.0001]	12.3452 [< 0.0001]
Auditor Type	−0.00151 [0.6378]	−0.04730 [< 0.0001]	−0.0145 [0.1179]	−1.2009 [0.0011]
n	18,550	19,021	19,958	17,613

(*continued on next page*)

TABLE 3 (continued)

Panel C: Raw Measures

Variable	**Discretionary Accruals**	**Current Accruals**
Intercept	−0.03388 [< 0.0001]	0.03202 [0.2360]
Tenure	0.00041 [0.0002]	−0.00032 [0.3792]
Age	0.00222 [< 0.0001]	0.00033 [0.4406]
Size	0.00361 [< 0.0001]	0.00318 [< 0.0001]
Industry Growth	0.00306 [0.1555]	0.00767 [0.2828]
Cash Flow	−0.36939 [< 0.0001]	−0.40922 [< 0.0001]
Auditor Type	−0.00688 [< 0.0001]	−0.00450 [0.2852]
n	37,571	37,571

Each cell contains a coefficient estimate and p-values in square brackets.
Omitting observations where the auditor-client relationship lasts for less than five years, observations where firms reported a merger or acquisition in the year, and the first year in the auditor-client relationship.

Accruals = one of the accruals measures presented in Figure 1;
Tenure = the number of consecutive years that the firm has retained the auditor;
Age = the number of years of data available on Compustat since 1980;
Size = the log of total assets;
Industry Growth = $\sum_{i=1}^{N} Sales_{i,t} / \sum_{i=1}^{N} Sales_{i,t-1}$ by two-digit SIC code;
Cash Flow = the firm's cash flow from operations divided by average total assets;
Auditor Type = an indicator set to 1 if the auditor is a Big N auditor, and set to 0 otherwise;
Industry = a control based on two-digit SIC codes (62 industries); and
Year = a control based on calendar year.

sales in the industry (i.e., is the lead auditor) or audits sales within 15 percent of the lead auditor. Our analyses reveal no significant differences (in the association between tenure and earnings quality) for specialist versus non-specialist auditors. Thus, our findings on specialization are not consistent with claims made by opponents of mandatory auditor rotation.

Next, to test whether period-specific characteristics around the time of the auditor change are driving our results, we eliminate firm-years representing the last year with the predecessor auditor. Again, our results are robust to this change.

In analysis not reported, we replicate all of our tests without requiring that the auditor audits the client in question for at least five years. The results obtained using this larger sample are generally consistent with those reported but are stronger. Any difference exists because the levels and dispersion of accruals are smaller for the long-tenure group (those where the relationship lasts for at least five years) than for the quick-turnover group (those where the relationship does not last for at least five years) for most measures of accruals. This magnifies our result on the tenure variable, but does not allow us to assess the relation between tenure and accruals as cleanly as does our chosen design.

To ensure that legitimate asset write-downs and impairment charges required by an incoming auditor in the initial year of the engagement do not explain the association between tenure and income-decreasing accruals, we eliminated the first year of an auditor-client relationship. As a sensitivity test, we also estimate Discretionary Accruals after eliminating Special Items. The results are consistent with the results reported in Tables 2 and 3.

Finally, the results in Panel B of Tables 2 and 3 are the result of estimating Model (1) using a truncated regression approach. In analysis not reported, we also estimate the model using ordinary least squares. The results are similar to those presented in the tables but confirm the bias due to truncation above (below) zero.

IV. CONCLUSIONS AND IMPLICATIONS

In this study, we examine the cross-sectional association between auditor tenure and audit quality. We use two measures of earnings quality to proxy for audit quality. These measures of earnings quality are drawn from prior literature that examines how accruals vary around events such as SEC violations and qualified opinions. These studies find that the level of accruals around the examined events is generally higher than in the absence of these events. We investigate the extent to which auditor tenure is associated with the dispersion in accruals and whether the recognition of income-increasing or income-decreasing accruals varies with auditor tenure. We contribute to the literatures on earnings quality and audit quality by documenting that longer auditor tenure is generally associated with less dispersion in the distributions of Discretionary Accruals and Current Accruals and that auditors appear to place greater constraints on both income-increasing and income-decreasing accruals as the auditor-client relationship lengthens.

Our study also provides useful evidence to regulators, members of the accounting profession, and financial statement users as they deliberate on the costs and benefits of mandatory auditor rotation. Specifically, our results suggest that under the current system, increased auditor tenure does not lead to reduced audit and earnings quality. If deteriorating earnings and/or audit quality are the motivation for calls for mandatory rotation, then our results do not support such an argument. However, our results also do not imply that forcing firms to remain with the same auditors would improve earnings quality or audit quality. Rather, Johnson and Lys (1990) show that allowing auditor turnover can improve audit quality because clients may switch to a more efficient provider (i.e., a new auditor) as their needs change. In this case, requiring a client to retain an auditor who is unable to provide adequate services could reduce audit quality. Further, we do not provide evidence on whether mandatory auditor rotation would improve or hamper audit quality. We merely provide evidence that under the current system of voluntary auditor rotation, audit quality does not appear to deteriorate with tenure.

Finally, our study does not address all instances of earnings management and our tests do not condition on managers' incentives to manage earnings. Therefore, additional avenues exist to provide evidence on the affects of extended auditor tenure on the quality of financial reporting.[21]

---

[21] For example, are restatements and Enforcement Actions (which have been interpreted as evidence of previous audit failures) more or less likely with longer audit tenure? Does audit tenure affect management's ability to manage earnings to avoid debt covenant violations, or to avoid losses, earnings declines, or missing analyst earnings forecasts?

REFERENCES

American Institute of Certified Public Accountants (AICPA). 1978. *The Commission on Auditors Responsibilities: Report, Conclusions and Recommendations*. New York, NY: AICPA.

———. 1992. *Statement of Position Regarding Mandatory Rotation of Audit Firms of Publicly Held Companies*. New York, NY: AICPA.

Anthony, J. H., and K. Ramesh. 1992. Association between accounting performance measures and stock prices: A test of the life cycle hypothesis. *Journal of Accounting and Economics* 15 (June/September): 203–227.

Barth, M. E., D. P. Cram, and K. K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (January): 27–58.

Bartov, E., F. A. Gul, and J. S. L. Tsui. 2000. Discretionary-accruals models and audit qualifications. *Journal of Accounting and Economics* 30 (December): 421–452.

Becker, C. L., M. L. DeFond, J. J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (Spring): 1–24.

Berton, L. 1991. GAO weighs auditing plan for big banks. *Wall Street Journal* (March 27): A3.

Collins, D. W., and P. Hribar. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (March): 105–134.

Davis, L. R., B. Soo, and G. Trompeter. 2002. Auditor tenure, auditor independence, and earnings management. Working paper, Boston College, Chestnut Hill, MA.

DeAngelo, L. E. 1981. Auditor independence, "low-balling" and disclosure regulation. *Journal of Accounting and Economics* 3 (August): 113–127.

———. 1986. Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders. *The Accounting Review* 61 (July): 400–420.

Dechow, P. M. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18 (July): 3–42.

———, and D. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (June): 235–250.

———, S. A. Richardson, and I. Tuna. 2001. Are benchmark beaters doing anything wrong? Working paper, University of Michigan, Ann Arbor, MI.

———, and I. D. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 35–59.

DeFond, M. L., and C. W. Park. 1997. Smoothing income in anticipation of future earnings. *Journal of Accounting and Economics* 32 (July): 115–139.

———, and K. R. Subramanyam. 1998. Auditor changes and discretionary accruals. *Journal of Accounting and Economics* 25 (February): 35–67.

Deis, D. R., and G. A. Giroux. 1992. Determinants of audit quality in the public sector. *The Accounting Review* 67 (July): 462–479.

———, and ———. 1996. The effect of auditor changes on audit fees, audit hours, and audit quality. *Journal of Accounting and Public Policy* 15 (Spring): 55–76.

Dunham, K. J. 2002. Firms that want to switch auditors find it takes time, money and faith. *Wall Street Journal* (March 15).

Ettredge, M., and R. Greenberg. 1990. Determinants of fee cutting on initial audit engagements. *Journal of Accounting Research* 28 (Spring): 198–210.

Francis, J. R., and J. Krishnan. 1999. Accounting accruals and auditor reporting conservatism. *Contemporary Accounting Research* 16 (Spring): 135–165.

———, E. L. Maydew, and H. C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18 (Fall): 17–34.

Frankel, R. M., M. F. Johnson, and K. K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings quality. *The Accounting Review* 77 (Supplement): 71–105.

LIBRARY

Geiger, M., and K. Raghunandan. 2002. Auditor tenure and audit quality. *Auditing: A Journal of Practice & Theory* 21 (March): 187–196.

Greene, W. H. 2000. *Econometric Analysis.* Englewood Cliffs, NJ: Prentice Hall.

Healy, P. M. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (April): 85–107.

———, and J. A. Wahlen. 1999. A review of the earnings management literature and it implications for standard setting. *Accounting Horizons* 13 (December): 365–383.

Heninger, W. G. 2001. The association between auditor litigation and abnormal accruals. *The Accounting Review* 76 (January): 111–126.

Johnson, W. B., and T. Lys. 1990. The market for audit services: Evidence from voluntary auditor changes. *Journal of Accounting and Economics* 12 (January): 281–308.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (Autumn): 193–228.

Lee, C. J., and Z. Gu. 1998. Low balling, legal liability and auditor independence. *The Accounting Review* 73 (October): 533–555.

Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, NY, September 28. Available at: http://www.sec.gov/news/speech/speecharchive/1998/spch220.txt.

———. 2000. Renewing the covenant with investors. Remarks delivered at the NYU Center for Law and Business, NY, May 10. Available at: http://www.sec.gov/news/speech/spch370.htm.

Maddala, G. S. 1997. *Introduction to Econometrics.* New York, NY: MacMillan Publishing Company.

Mautz, R. K., and H. A. Sharaf. 1961. *The Philosophy of Auditing.* American Accounting Association Monograph No. 6. Sarasota, FL: American Accounting Association.

Palmrose, Z-V. 1986a. Audit fees and auditor size: Further evidence. *Journal of Accounting Research* 24 (Spring): 97–110.

———. 1986b. Litigation and independent auditors: The role of business failures and management fraud. *Auditing: A Journal of Practice & Theory* (Fall): 90–113.

———. 1991. Trials of legal disputes involving independent auditors: Some empirical evidence. *Journal of Accounting Research* (Supplement): 149–185.

Petty, R., and S. Cuganesan. 1996. Auditor rotation: Framing the debate. *Australian Accountant* 66 (May): 40–41.

PricewaterhouseCoopers LLP. 2002. *Mandatory Rotation of Audit Firms: Will It Improve Audit Quality?* New York, NY: PricewaterhouseCoopers LLP.

Richardson, S., R. G. Sloan, M. T. Soliman, and I. Tuna. 2002. Information in accruals about the quality of earnings. Working paper, University of Michigan, Ann Arbor, MI.

Securities Exchange Commission (SEC). 1994. *Staff Report on Auditor Independence.* Washington, D.C.: Government Printing Office.

Simon, D. T., and J. R. Francis. 1988. The effects of auditor change on audit fees: Tests of price cutting and price recovery. *The Accounting Review* 63 (April): 255–269.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (July): 289–315.

Teoh, S. H., I. Welch, and T. J. Wong. 1998a. Earnings management and the long-run market performance of initial public offerings. *The Journal of Finance* 53 (December): 1935–1974.

———, T. J. Wong, and G. Rao. 1998b. Are accruals during initial public offerings opportunistic? *Review of Accounting Studies* 3: 175–208.

Thomas, J. K., and H. Zhang. 2001. Inventory changes and future returns. Working paper, Columbia Business School, New York, NY.

U.S. Senate. 1976. *The Accounting Establishment.* Prepared by the Subcommittee on Reports, Accounting and Management of the Committee on Governmental Affairs. Washington, D.C.: Government Printing Office.

———. 1977. *Improving the Accountability of Publicly Owned Corporations and Their Auditors.* Prepared by the Subcommittee on Reports, Accounting and Management of the Committee on Governmental Affairs. Washington, D.C.: Government Printing Office.
Wallman, S. 1996. The future of accounting, Part III: Reliability and auditor independence. *Accounting Horizons* 10 (December): 76–97.
Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76 (July): 357–373.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 801–818

Effects of Authoritative Guidance Availability and Audit Committee Effectiveness on Auditors' Judgments in an Auditor-Client Negotiation Context

Terence Bu-Peow Ng
Hun-Tong Tan
Nanyang Technological University

ABSTRACT: In this paper, we present the results of an experiment that investigates the effects of two contextual features—the availability of authoritative guidance and the effectiveness of the client's audit committee—on auditors' perceived outcome of auditor-client negotiations concerning an audit adjustment that affects the client's ability to meet analysts' forecasts. Results show that auditors' perceived negotiation outcome is jointly influenced by authoritative guidance availability and audit committee effectiveness. Specifically, authoritative guidance availability has a greater effect on auditors' perceived negotiation outcome in the absence of an effective audit committee than in its presence. We also find that concessionary moves by the client increase auditors' propensity to concede to the client's preferred position.

Keywords: *authoritative guidance; audit committee; audit adjustments; negotiation.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

The audited financial statement is the outcome of a negotiation process between the auditor and the client (e.g., Antle and Nalebuff 1991). In this study, we investigate the joint effects of two contextual features present in the audit environment—availability of authoritative guidance and the effectiveness of the client's audit committee—on auditors' perceived negotiated outcomes within an auditor-client negotiation context. These factors are relevant to current discussions on principles-based versus rules-based accounting standards (Anonymous 2002), and to calls for greater audit committee involvement (SAS No. 89 [AICPA 1999a] and SAS No. 90 [AICPA 1999b]). Further, while both factors likely contemporaneously co-exist in a given setting, evidence on their *joint* effects is lacking, and existing research provides no evidence on whether the effect of one factor is moderated by the presence of the other. By examining how these two factors

---

We appreciate the support of the participating auditors, and, especially, the help of the firm's Assurance Services Director and his assistant. We thank Mark Nelson (special editor) and two anonymous reviewers for their insightful comments. We also thank Amit Das, Mike Gibbins, Jun Han, Jane Kennedy, Bill Messier, Steve Salterio, and NTU colleagues for helpful comments, and Magdalene Lim for programming assistance.

Editor's note: This paper was accepted by Mark W. Nelson, Special Editor.

Submitted August 2002
Accepted February 2003

interact, we identify boundary conditions under which prior research findings hold. Specifically, our study suggests that the influence of authoritative guidance on auditor judgment and financial reporting quality (e.g., Salterio and Koonce 1997; Nelson et al. 2002) is contingent on audit committee effectiveness, and that the influence of audit committee strength on auditor effectiveness and financial reporting quality (e.g., Carcello and Neal 2000) depends on the availability of authoritative guidance.

We conduct an experiment in which auditors provide judgments relating to a proposed audit adjustment that is quantitatively immaterial but that will affect the client's ability to meet or beat the analysts' consensus forecast. We manipulate availability of precise authoritative guidance and audit committee effectiveness as between-subject variables. This setting allows us to gain some insights into issues related to audit adjustments and audit committee communication that are the subject matter of authoritative pronouncements (SAS No. 89 and SAS No. 90). In addition, we use an audit adjustment context that relates to the potential inappropriate choice of revenue-recognition method because this is a major avenue for earnings management attempts (Committee of Sponsoring Organizations [COSO] 1999; Nelson et al. 2002) that has prompted recent issuance of authoritative guidance on revenue recognition by the AICPA (Anonymous 1999) and the SEC (Staff Accounting Bulletin No. 101; SEC 1999a). More generally, in a post-SAS No. 90 environment, auditors and audit committee members are likely to be concerned about audit adjustments pertaining to accounting method choice that affect financial reporting quality.

We find that authoritative guidance availability has a greater effect on auditors' perceived negotiation outcome when the client's audit committee is ineffective than when it is effective. Similarly, audit committee effectiveness has a greater effect on auditors' perceived negotiation outcome in the absence of authoritative guidance than in its presence. In additional analyses, we provide some evidence that auditors' propensity to concede to the client's preferred position is also influenced by these two factors, and by concessionary moves made by the client. Our results suggest that precise standards and effective audit committees are potential substitutes/compensating mechanisms for enhancing auditor effectiveness and financial reporting quality. Specifically, our results suggest that corporate governance mechanisms such as effective audit committees may mitigate the potential adverse effects of imprecise accounting rules by bolstering auditors' positions during negotiations with the client, and provide support for continual efforts to enhance the effectiveness of audit committees (e.g., Blue Ribbon Committee 1999; National Association of Corporate Directors 1999).

The following section develops our research hypotheses. This is followed by a description of the research methods used and presentation of the results. The paper concludes with a discussion of the study's main findings, limitations, and opportunities for future research.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

Pronouncements issued by standard setters and regulatory bodies (such as the FASB, ASB, and SEC) provide authoritative guidance on the treatment and resolution of potentially ambiguous accounting issues. The availability of precise authoritative guidance that indicates the (more) appropriate accounting treatment among available alternatives provides justification for auditors' choice of accounting methods and treatment of accounting issues (Kennedy et al. 1997). Prior literature has documented that auditors respond to incentives to permit aggressive reporting by interpreting imprecise standards or precedents as allowing such reporting (Hackenbrack and Nelson 1996; Salterio and Koonce 1997; Nelson et al. 2002). Thus, besides serving a technical justification role, precise authoritative guidance can also be viewed as part of the repertoire of bargaining "ammunition" available to auditors in their negotiations with clients on aggressive accounting choices. Such authoritative guidance generally recommends or mandates how accounting issues are to be treated, and therefore limits accounting choices available to the client (e.g., see Libby et al. [2002] for a review of this literature).

Like authoritative guidance availability, audit committees form part of the institutional environment that auditors and clients face, and are an important corporate governance mechanism with a direct role in the oversight of the financial reporting process (Nichols and Price 1976; Johnstone et al. 2001). Unlike authoritative guidance availability, audit committees do not play a technical justification role. The effectiveness of an audit committee has little direct bearing on the technical justifiability of an accounting choice. However, similar to authoritative guidance, effective audit committees also help to bolster auditors' bargaining power in their negotiations with clients, because they signal that auditors' judgments with respect to appropriate accounting will be considered carefully, even if they could reflect unfavorably on management. Recently issued standards requiring communication between auditors and audit committees on unrecorded misstatements (SAS No. 89) and on the quality of accounting principles/estimates (SAS No. 90) enhance the importance of the audit committee as a party that influences the outcome of accounting negotiations with the client.

Despite the influence that both authoritative guidance and effective audit committees likely have on auditor effectiveness and financial reporting quality, evidence on their interactive effects is lacking. For example, Trompeter (1994) finds that auditors make more aggressive recommendations when GAAP has fewer limits on the range of acceptable accounting alternatives. Other studies show that auditors have a tendency to interpret vague standards as indicating support for management's aggressive or conservative financial reporting, depending on their incentives (Hackenbrack and Nelson 1996). However, these studies do not examine the influence of clients' corporate governance mechanisms, so it is unclear whether the findings hold even when the client has an effective corporate governance structure (e.g., an effective audit committee).[1] Another stream of research investigates the effect of audit committee existence/effectiveness on financial reporting quality (DeFond and Jiambalvo 1991; Dechow et al. 1996; McMullen 1996; Beasley et al. 2000), and on auditor effectiveness (Carcello and Neal 2000; Cohen and Hanno 2000). However, these studies do not examine how the impact of audit committee effectiveness on auditors' judgments may vary depending on the level of authoritative guidance available.

Consider a situation where auditors are told that management is reluctant, even after several rounds of discussions and negotiations, to record an audit adjustment relating to a revenue-recognition issue that will result in their inability to meet the analysts' consensus forecast. Suppose auditors are asked to indicate the amount of the audit adjustment that will ultimately be recorded in the audited financial statements. As discussed earlier, auditors' judgments are likely influenced by the availability of authoritative guidance on the underlying technical issue (here, revenue recognition). Furthermore, given the background information, it should be apparent to the auditors that the amount ultimately recorded is the result of a negotiation process with the client's management, and consideration of bargaining power should factor into the auditors' assessments of the ultimate negotiated outcome. In particular, besides authoritative guidance availability, auditors are also likely to consider the effectiveness of the client's audit committee because the ultimate negotiated audit adjustment likely occurs after meetings with the audit committee to discuss audit findings (see SAS No. 89 and SAS No. 90).

---

[1] An earlier study by Knapp (1987) finds that audit committee members are more likely to support auditors' stance in an auditor-client conflict where objective technical standards also support the auditors' stance (i.e., when the issue relates to disclosure of subsequent events), but not otherwise (i.e., when the issue relates to the materiality of an audit adjustment). Knapp's (1987) study focuses on the judgments of audit committee members, not auditors, and is conducted before the introduction of several regulatory measures that enhance the audit committee's role in the financial reporting process (e.g., SAS No. 89, SAS No. 90).

The availability (or the lack) of authoritative guidance can be viewed as an external (technical) constraint on the eventual audit adjustment ultimately recorded, but the effect of this constraint may be moderated by the effectiveness of the audit committee (a source of bargaining power). Negotiation research using buyer-seller contexts suggests that bargaining power can moderate the effect of external constraints on negotiation outcomes (see McAlister et al. [1986], which investigates the context of an externally set profitability constraint). Authoritative guidance availability and audit committee effectiveness can also be viewed as alternative sources of bargaining power for auditors in their negotiations with clients (e.g., Nichols and Price 1976; Gibbins et al. 2001), or as compensating control mechanisms over financial reporting quality (Johnstone et al. 2001; see also Brown and Solomon [1990] on auditors' configural processing of internal controls). For these reasons, authoritative guidance availability and audit committee effectiveness likely interact to influence auditors' perceived negotiation outcome. Specifically, we predict that in the absence (presence) of an effective audit committee, auditors are less (more) likely to perceive that they are able to negotiate successfully with client's management for a more conceptually sound accounting method that will cause the client to miss the analysts' consensus forecast. Thus, the availability of authoritative guidance (which serves as a technical constraint on auditors' decisions, as well as an alternative source of bargaining power) is likely to exert a greater influence on auditors' perceived negotiation outcome (as proxied by whether the negotiated adjustment perceived by auditors to be ultimately recorded will cause the client to miss the analysts' forecasts) in the absence of an effective audit committee than in its presence. Our hypothesis on the interactive effect of authoritative guidance availability and audit committee effectiveness on auditors' perceived negotiation outcome is formally stated below:

> **H1a:** The proportion of auditors who believe that the amount of the audit adjustment ultimately recorded will result in the client's failure to meet the analysts' consensus forecast is more positively affected by the availability of authoritative guidance when the audit committee is ineffective than when the audit committee is effective.

An alternative perspective on the interaction effect of authoritative guidance availability and audit committee effectiveness is that audit committee effectiveness is more likely to affect auditors' perceived negotiation outcome when authoritative guidance is absent than when it is present. The absence of authoritative guidance increases the level of ambiguity, which could increase the role and impact of bargaining power arising from differential audit committee effectiveness. The joint effect of authoritative guidance availability and audit committee effectiveness on auditors' perceived negotiation outcome is restated as follows:

> **H1b:** The proportion of auditors who believe that the amount of the audit adjustment ultimately recorded will result in the client's failure to meet the analysts' consensus forecast is more positively affected by audit committee effectiveness when authoritative guidance is absent than when authoritative guidance is present.

III. METHOD

Design

We employ a 2 × 2 between-subject factorial design, with *Guidance Availability* (absent, present) and *Audit Committee Effectiveness* (low, high) as the independent variables. Case information provided to participants contained a brief description of the hypothetical company, Libra Inc., selected financial information, and the accounting issue involving an audit adjustment proposed by the audit staff. The audit adjustment represents the difference in the revenue recognized from a new sales contract between the method used by the client and that proposed by the audit staff. It amounts to 3 cents per share, which is quantitatively immaterial (i.e., less than 1 percent of sales, earnings, and total assets). However, since the EPS before the audit adjustment is $3.68 and the analysts' consensus forecast is $3.66, booking the full amount of the audit adjustment will result in a failure by the client to meet the analysts' consensus forecast by 1 cent per share, a qualitative factor considered to be relevant for materiality considerations under Staff Accounting Bulletin No. 99 (SEC 1999b). Given that it has been argued that missing analysts' consensus forecasts by even 1 cent per share is material to financial statement users (e.g., Fox 1997; Levitt 1998; Brown 2001), we use 1 cent per share as the amount of the analysts' consensus forecast missed to provide a boundary condition for testing the hypothesized effects.

The case information indicates that the amount of the audit adjustment represents about 0.08 percent of sales, 0.8 percent of earnings, and 0.09 percent of total assets, and that booking the full amount of the audit adjustment will cause the company's adjusted EPS to fall short of the analysts' consensus forecast by 1 cent per share. We instructed all participants, before they made their recommendations on the audit adjustment, that the client's management had strong reservations about the proposed audit adjustment and were not willing to record any amount of the audit adjustment that would reduce the revenue recognized.

The first independent variable manipulates the availability of authoritative guidance on the accounting issue. We provided general descriptions of the accounting issue (without reference to specific revenue-recognition methods) to ensure that auditors made judgments based on their average experience of similar auditor-client negotiation situations, which better controls for differential technical knowledge about the accounting issue among auditors (e.g., see Libby and Kinney 2000). In all conditions, we informed participants that "there are some reasons to believe that the method proposed by the audit staff may be more conceptually sound (relative to the method used by the client)." This ensures that there are grounds to believe that the auditors' proposed method is the more appropriate method regardless of authoritative guidance availability, and provides a basis for the audit adjustment to be proposed and supported in the first place. In particular, SAB No. 101 notes that "in the absence of authoritative literature...the existing authoritative accounting standards as well as the broad revenue-recognition criteria specified in the FASB's conceptual framework that contain basic guidelines for revenue recognition" should be considered. Thus, according to SAB No. 101, the presence of authoritative guidance is not a necessary condition for the determination of the appropriate accounting treatment applicable for specific transactions and situations, as existing conceptual frameworks are to be used in its absence. We informed participants in the *No Guidance* condition that "there is no authoritative guidance indicating which of the two methods is the more appropriate method under GAAP," and those in the *Guidance* condition that "there is also authoritative guidance indicating that the method proposed by the audit staff is the more appropriate method under GAAP."

The second independent variable manipulates the effectiveness of the client's audit committee (in terms of its independence from management). We informed participants in the *Ineffective Audit Committee* condition that "in your past dealings with the audit committee, you have found that it has a tendency to be sympathetic toward the management's position," and those in the *Effective Audit Committee* condition that "in your past dealings with the audit committee, you have found that it consistently exercises independent judgment without a tendency to be sympathetic toward the management's position."

After reading the case information, participants provided two judgments: their recommendations on the amount of the audit adjustment to be booked, and assessments of the perceived negotiation outcome. At the first stage of the experiment, participants were posed the following question: "Suppose that you are about to meet Libra's management to discuss the audit adjustment. What is the amount of the proposed audit adjustment that you would recommend to be booked?" We asked participants to indicate an amount from \$0 (i.e., completely waive the audit adjustment) to \$300,000 (i.e., completely book the audit adjustment), and to indicate the rationale for their responses. After answering the above questions, they proceeded to the second stage of the experiment where we provided them the following information: "Suppose that after a few rounds of discussions and negotiations, Libra's management has reiterated its strong reservations about the proposed audit adjustment, but has indicated that it is now willing to record up to \$100,000 of the proposed audit adjustment (i.e., up to 1 cent per share), but no more than that." Management's willingness to record up to \$100,000 is held constant across all experimental conditions.[2] Following the approach used in Libby and Kinney (2000), we elicited auditors' estimates of the auditor-client negotiation outcome (which form our main dependent variable) by asking this question: "In light of the client's stand above, what is the amount of the proposed audit adjustment that you believe will ultimately be recorded in Libra's audited financial statements?"[3]

Upon providing the above judgments relating to the audit adjustment, the participants responded to debriefing questions.

Participants and Administration

Participants were audit managers from various U.S. offices of a Big 4 firm.[4] The experiment was administered online with the help of a contact person from the firm's Assurance Services Division. Through the contact person, participants were randomly assigned and emailed the different versions of the instrument corresponding to the four cells of our 2 × 2 design. The email was accompanied by a cover message from the Assurance Services Director requesting participation in the study and assuring participants of the confidentiality of their responses. In our instructions, we reminded participants to perform the judgment tasks independently and to provide their email address.[5]

---

[2] Prior research suggests that auditors are open to booking partial amounts of audit adjustments (e.g., Libby and Kinney 2000; Braun 2001). In the context of our study, it is possible that changes in parameters and estimates within the accounting methods may be used to justify partial audit adjustments.

[3] Libby and Kinney (2000, 387) argue that auditors "provide relatively informed and unbiased estimates of reported (audited) earnings" resulting from auditor-client negotiations (relative to management) given that they "are in the best position to estimate their own behavior" and "are likely to have experience negotiating corrections across several client contexts."

[4] Managers are appropriate participants for our study because they likely provide inputs to partners in the latter's negotiations with clients' management, and should also be aware of negotiation outcomes. They also may directly engage in negotiations with clients' management on some audit adjustments. Supporting the usefulness of audit managers' data, Nelson et al. (2002) find no differences in audit managers' and partners' responses describing audit outcomes with respect to clients' earnings management attempts.

The data-collection program used allows participants to proceed from one stage to another only after they have responded to questions in the preceding stages. The program has features that limit respondents from scrolling back to earlier screens to change their responses.[6] As an added control, we designed the program to immediately email completed responses at each stage to the researchers. We pilot-tested the instrument with five faculty members who had between four to eight years of audit experience, and modified it based on the feedback received from the Assurance Services Director and two field managers from the participating firm. We obtained responses from 113 out of 224 audit managers contacted via email, yielding a response rate of 50 percent. We eliminated responses from 12 participating auditors with incomplete or non-interpretable answers. The remaining 101 auditors had an average of 8.3 years of audit experience (standard deviation = 3.1 years), 38 percent of which was spent with SEC-listed clients.

IV. RESULTS

Manipulation and Other Checks

We ask several questions to assess whether auditors correctly interpret our task and manipulations. A large majority do so, with the following exceptions: three auditors who indicate that they are not aware that booking the full amount of the audit adjustment would cause a failure to meet the analysts' consensus forecast; two auditors who indicate that authoritative guidance is available when it is not, and vice versa; and seven auditors in the effective (ineffective) audit committee condition who rate the audit committee below (above) the mid-point of the audit committee effectiveness scale (0 = not effective at all; 10 = extremely effective).[7] We eliminate the responses of these auditors from further analyses.[8] Of the remaining 89 auditors, 50 respond to the versions corresponding to the four experimental treatments related to our main experiment, while 39 respond to additional treatments used for further analysis (see discussion under "Additional Analyses" in this section).

We also ask auditors additional questions, related to the task, on 11-point scales with the appropriate verbal anchors (e.g., 0 = not likely at all; 10 = extremely likely). Auditors express a strong belief that management's reservations about the proposed audit adjustment are likely due to their desire to meet the analysts' consensus forecast (mean = 8.61, standard deviation = 1.63), suggesting that they are aware of management's incentive to avoid recording the audit adjustment. On average, auditors rate the materiality of the audit adjustment to be moderate and not low (mean = 6.07; standard deviation = 3.16), suggesting that they do consider the analysts' forecast factor in their materiality assessments. They rate the client's engagement risk (defined as the likelihood that the audit firm will suffer a loss as a result of its association with the client) to be moderate (mean = 4.96; standard deviation = 2.21). Auditors' materiality and engagement risk assessments do not vary systematically across experimental conditions ($p > 0.741$).[9] Across experimental conditions, auditors rate the client's management to be of moderate integrity (mean = 5.19, standard deviation = 1.85). We also ask the auditors to rate separately the importance of authoritative guidance and audit committee effectiveness in mitigating biased financial reporting by clients in general. The mean rating for the importance of authoritative guidance is significantly higher than that for audit committee effectiveness (means = 8.61 and 7.85, respectively; $p = 0.004$), suggesting that auditors across conditions believe that authoritative guidance has a major influence on financial reporting quality.

---

[5] By asking participants for their email addresses, we intend that participants become aware that, even though their responses are confidential, the researchers may potentially contact them. This should induce greater accountability than if their responses were anonymous. Only three participants did not provide an email address. Eliminating these participants from the analyses does not alter the results of the hypothesis tests.

[6] Participants who double-click on the "back" button in quick succession can return to the earlier screen (a feature inherent in the program). However, the program is designed such that we are alerted if they change their responses. Only two participants changed their earlier responses. These responses are treated as uninterpretable, and have been eliminated from the analyses.

[7] The average ratings in the *Effective Audit Committee* and *Ineffective Audit Committee* conditions are 6.82 and 2.93, respectively ($p = 0.000$). Unless otherwise stated, two-tailed p-values are reported. One-tailed p-values are reported for directional tests.

[8] These questions screen out auditors who may not have carefully read the contents of our Internet-administered research instrument. Auditors who misinterpret the experimental manipulations add noise to the analyses and weaken the power of the tests. The manipulation check failures occur evenly across the four main experimental and three additional conditions. Including these responses in the analyses yield similar results for H1a (where the effect of authoritative guidance availability is only statistically significant when the audit committee is ineffective; $p = 0.028$, one-tailed) but not for H1b (where the effect of audit committee effectiveness is not significant for both guidance conditions; $p > 0.271$).

Tests of Hypotheses

The dependent variable we use in the hypothesis tests is a dichotomous variable coded based on whether the amounts judged by the auditors to be ultimately booked after negotiations with the client are above $200,000, an amount that will result in a failure by the client to meet the analysts' consensus forecast (coded as 1), or otherwise (coded as 0). This is because what matters to the client and users of audited financial statements (in the context of this study) is not the magnitude of the audit adjustment booked, per se (given that the amount of the audit adjustment is quantitatively immaterial), but whether the amount booked will enable the client to meet or beat the analysts' consensus forecast. Booking an amount of $200,000 (i.e., 2 cents per share) reduces the adjusted EPS to $3.66, which would enable the client to just meet the analysts' consensus forecast. Thus, the dichotomous variable distinguishes between judgments that are consistent with the client's preference and those that are not.[10]

Table 1, Panel A presents the descriptive statistics on the proportion of auditors who believe that the audit adjustment ultimately booked after auditor-client negotiation will result in the client's failure to meet the analysts' consensus forecast. The mean and standard deviation of the actual amounts of the adjustment perceived to be ultimately booked by auditors are also presented in Panel A. On average, across all experimental conditions, only an average of 26 percent believe that the amount ultimately booked after negotiations will result in the client's failure to meet the analysts' consensus forecast.

Panel B of Table 1 reports the categorical analysis results (using the likelihood ratio tests of partial associations of the log-linear model) relating to auditors' perceived negotiation outcome. The main effect of *Guidance Availability* is statistically significant ($p = 0.031$), but not that of *Audit Committee Effectiveness* ($p = 0.309$). Of relevance to the tests of H1a and H1b is the statistically significant interaction effect between *Guidance Availability* and *Audit Committee Effectiveness* ($p = 0.030$, one-tailed). H1a predicts that

---

[9] Including auditors' materiality or engagement risk assessments as an independent variable (coded as a dichotomous variable based on median splits) in the hypothesis tests yield similar results. While the main effects of materiality and engagement risk assessments are statistically significant ($p < 0.021$), these assessments do not interact significantly with the key independent variables of interest.

[10] Coding the dichotomous variable based on whether the amounts are above $250,000 yields similar results for the tests of hypotheses. We also obtain similar results when the hypotheses are tested based on the amounts reported by auditors instead of the dichotomous variables.

TABLE 1
Effects of Authoritative Guidance Availability and Audit Committee Effectiveness on Auditors' Perceived Negotiation Outcome

Panel A: Descriptive Statistics

			Audit Committee Effectiveness		
			Ineffective	**Effective**	**Average**
Authoritative Guidance	Absent	Proportion[a]	0.00	0.25	0.13
		Mean (s.d.) Adjustment[b]	50 (52)	117 (127)	83 (101)
		Sample size	12	12	24
	Present	Proportion[a]	0.38	0.38	0.38
		Mean (s.d.) Adjustment[b]	162 (126)	154 (127)	158 (124)
		Sample size	13	13	26
	Average	Proportion[a]	0.20	0.32	0.26
		Mean (s.d.) Adjustment[b]	108 (112)	136 (125)	122 (118)
		Sample size	25	25	50

Panel B: Tests of Partial Associations of the Log-Linear Model: Effects on Auditors' Perceived Negotiation Outcome

Effect	df	Likelihood Ratio Chi-Square	p-value
Guidance Availability	1	4.666	0.031
Audit Committee Effectiveness	1	1.033	0.309
Guidance Availability × Audit Committee Effectiveness	1	3.556	0.030*

---

* Denotes one-tailed p-value.

[a] Proportion = proportion of auditors who believe that the amount of the audit adjustment ultimately recorded will result in the client's failure to meet the analysts' consensus forecast.

[b] Mean (s.d.) Adjustment = mean (standard deviation of) audit adjustment (in $000) that auditors believe will ultimately be recorded in the client's audited financial statements.

the effect of authoritative guidance on auditors' perceived negotiation outcome is moderated by the effectiveness of the audit committee. Tests of simple main effects indicate that, when the audit committee is ineffective, a significantly higher proportion of auditors believe that the amount ultimately booked will result in the client's failure to meet the analysts' consensus forecast when authoritative guidance is present than when it is absent (proportions = 0.38 and 0.00, respectively; $\chi^2 = 7.70$, $p = 0.003$, one-tailed). In contrast, when the audit committee is effective, the difference in proportions between the two guidance conditions is not statistically significant (proportions = 0.38 and 0.25, respectively; $\chi^2 = 0.52$, $p = 0.469$). These results are consistent with H1a.

Hypothesis 1b predicts that the effect of audit committee effectiveness on auditors' perceived negotiation outcome is moderated by the availability of authoritative guidance. Tests of simple main effects indicate that, when authoritative guidance is absent, a significantly higher proportion of auditors believe that the amount ultimately booked will result in the client's failure to meet the analysts' consensus forecast when the client's audit committee is effective than when it is ineffective (proportions = 0.25 and 0.00, respectively; $\chi^2 = 4.57$, $p = 0.016$, one-tailed). In contrast, when authoritative guidance is present, there is no difference in the proportions between the two audit committee effectiveness conditions (proportions = 0.38 and 0.38, respectively; $\chi^2 = 0.00$, $p = 1.00$). These results are consistent with H1b.

Additional Analyses

Auditors' Propensity to Concede to Client's Preferred Position

In this subsection, we assess whether auditors' perceived negotiation outcome deviates from their initial recommendations in a manner that suggests a concession toward the client's preferred position. To the extent that auditors' initial recommendations approximate their assessments of the most appropriate audit adjustment, a move away from their initial recommendations toward the client's preferred position indicates a shift in the auditors' (apparent) objectivity.

In the context of our study, auditors' initial recommendations likely approximate their true beliefs of the most appropriate audit adjustment. Although these initial recommendations can be viewed as opening moves that may be strategically "inflated" to incorporate consideration of bargaining power (e.g., Yukl 1974; Pruitt and Carnevale 1993), institutional constraints in the audit environment limit this possibility. In particular, auditors' initial recommendation likely is recorded as a proposed adjusting entry, and auditors may avoid making extreme recommendations if they consider it inappropriate or unprofessional to propose an incorrect amount under existing accounting or auditing standards and guidelines. We asked a subset of the auditors participating in additional experimental conditions (reported later; $n = 12$) to indicate on 11-point scales (0 = not at all; 10 = to a large extent) the degree to which their initial recommendations (a) are based on their assessments of the most appropriate amount that they think should be recorded in the client's financial statements, and (b) are higher than the most appropriate amounts that they think should be recorded (to incorporate an expectation of potential bargaining down by the client during subsequent negotiations). The mean responses for these two questions are 9.33 and 0.67, respectively, with the difference being highly significant ($p = 0.000$). These results, from a subset of the participants, indicate that auditors strongly believe that their initial recommendations are based on their assessments of the most appropriate amounts that should be recorded, and are not strategically biased upward to factor in anticipated adjustments likely to occur in subsequent negotiations.[11]

As in our main hypothesis, we anticipate that authoritative guidance availability and audit committee effectiveness will interact such that auditors' propensity to concede to the client's position is more significantly affected by the availability of authoritative guidance when the audit committee is ineffective than when the client's audit committee is effective.[12] We define propensity to concede as referring to a situation where auditors initially recommend booking an amount that will result in the client's failure to meet the analysts' consensus forecast, but perceive that such an amount will not be ultimately recorded after negotiations with the client. Thus, our analyses involve only the 34 auditors who initially recommend booking an amount that will result in the client's failure to meet the analysts' consensus forecast. Results are presented in Table 2. As an overall test, we conduct a categorical analysis with *Guidance Availability* and *Audit Committee Effectiveness* as independent variables and, as dependent variable, a dichotomous variable based on whether the audit adjustment perceived to be ultimately recorded in the audited financial statements results in the client's failure to meet the analysts' consensus forecast. Table 2, Panel B shows that, as expected, the two-way interaction is statistically significant (p = 0.025, one-tailed).[13]

---

[11] We conduct a categorical analysis with authoritative guidance availability and audit committee effectiveness as independent variables and, as dependent variable, a dichotomous variable coded based on whether auditors recommend, at the initial stage, booking an amount that will result in the client's failure to meet the analysts' consensus forecast. Consistent with the idea that auditors base their initial recommendations predominantly on the technical justifiability of the audit adjustment, we find a statistically significant main effect of guidance availability (mean proportions = 0.81 and 0.54 for the *Guidance* and *No Guidance* conditions, respectively; $p = 0.042$), and neither the main effect of audit committee effectiveness nor the interaction effect is statistically significant ($p > 0.527$).

[12] Gibbins and Salterio's (2002) model suggests that the absence (presence) of authoritative guidance, which affects auditors' issue assessments, increases the likelihood of overlap (non-overlap) between auditor- and client-desired outcomes, and consequently the tendency of auditors to concede (not concede) toward the client-preferred adjustment. Similarly, auditors' perceived bargaining power, which increases with audit committee effectiveness and authoritative guidance availability, decreases the tendency of auditors to concede toward the client-preferred adjustment. In the context of Gibbins and Salterio's (2002) model, we assume that, within limits, auditors' motivations are generally not adversarial to reaching client-desired outcomes.

TABLE 2
Effects of Authoritative Guidance Availability and Audit Committee Effectiveness on Auditors' Propensity to Concede

Panel A: Descriptive Statistics

			Audit Committee Effectiveness		
			Ineffective	**Effective**	**Average**
Authoritative Guidance	Absent	Proportion[a]	1.00	0.57	0.77
		Mean (s.d.) Adjustment[b]	233 (52)	114 (121)	169 (111)
		Sample size	6	7	13
	Present	Proportion[a]	0.50	0.55	0.52
		Mean (s.d.) Adjustment[b]	100 (115)	127 (127)	114 (120)
		Sample size	10	11	21
	Average	Proportion[a]	0.69	0.56	0.62
		Mean (s.d.) Adjustment[b]	150 (115)	122 (122)	135 (118)
		Sample size	16	18	34

Panel B: Tests of Partial Associations of the Log-Linear Model: Effects on Auditors' Propensity to Concede

Effect	**df**	**Likelihood Ratio Chi-Square**	**p-value**
Guidance Availability	1	2.200	0.138
Audit Committee Effectiveness	1	0.705	0.401
Guidance Availability × Audit Committee Effectiveness	1	3.823	0.025*

The results presented in this table involve only the 34 (out of 50) auditors who initially recommend booking an amount that will result in the client's failure to meet the analysts' consensus forecast.

* Denotes one-tailed p-value.

[a] Proportion = proportion of auditors who believe that the amount of the audit adjustment ultimately recorded will not result in the client's failure to meet the analysts' consensus forecast.

[b] Mean (s.d.) Adjustment = mean (standard deviation) of the difference (in $000) between the amount of the audit adjustment that auditors initially recommend to be booked and the amount the they believe will ultimately be recorded in the client's audited financial statements.

---

[13] We also analyze the difference in auditors' judgments between the two stages (initial recommendation and perceived ultimate outcome), without consideration of whether they do or do not accede to the client's position in their initial recommendations. To do this, we conduct a mixed categorical analysis, with *Stage* (initial recommendation, perceived ultimate outcome) as a within-subject variable and *Guidance Availability* and *Audit Committee Effectiveness* as between-subject variables. Contrary to the main findings, we detect no statistically significant three-way interaction (p = 0.379). This result is likely obtained because the analysis includes auditors who *do not* initially recommend booking an amount that will result in the client's failure to meet the analysts' consensus forecast and, therefore, are not considered to have conceded to the client at the second stage. Including these auditors in the analysis weakens the power of the test on the joint effects of guidance availability and audit committee effectiveness on auditors' propensity to concede to the client.

Results of McNemar tests (Siegel and Castellan 1988) within each treatment group indicate that the change in auditors' position among those who initially recommend booking an amount that will result in the client's failure to meet the analysts' consensus forecast is at least marginally significant for each of the treatment groups, with the proportion of auditors having the propensity to concede ($P_{concede}$) ranging from 0.50 to 1.00 ($p < 0.063$, one-tailed), as summarized in Table 2, Panel A. Tests of differences between treatment groups indicate that when the audit committee is ineffective, $P_{concede}$ is significantly higher when authoritative guidance is absent than when it is present (proportions = 1.00 and 0.50, respectively; $\chi^2 = 6.01$, $p = 0.007$, one-tailed). In contrast, when the audit committee is effective, the difference in $P_{concede}$ between the two guidance conditions is not statistically significant (proportions = 0.57 and 0.55, respectively; $\chi^2 = 0.01$, $p = 0.914$). Similarly, when authoritative guidance is absent, $P_{concede}$ is significantly higher when the client's audit committee is ineffective than when it is effective (proportions = 1.00 and 0.57, respectively; $\chi^2 = 4.49$, $p = 0.017$, one-tailed). In contrast, when authoritative guidance is present, there is no difference in $P_{concede}$ between the two audit committee effectiveness conditions (proportions = 0.50 and 0.55, respectively; $\chi^2 = 0.43$, $p = 0.835$). These results suggest that authoritative guidance availability and audit committee effectiveness jointly influence auditors' propensity to concede toward the client's position.

Effects of Client's Concessionary Moves during Negotiations

Prior research suggests that negotiation strategies can influence negotiated outcomes (e.g., Antle and Nalebuff 1991; Neale and Northcraft 1991). To examine the effects of concessionary moves by the client during negotiations on auditors' judgments, we include three additional treatments. Two-thirds of the data for these treatments are randomly distributed and were collected at the same time as those for the four main experimental treatments; the remaining one-third of the data was collected after the main data collection, and participants were directed to cells with the fewest observations.

We expect that auditors will have a greater propensity to concede to the client's preferred position in the presence of concessionary moves by the client (e.g., see Yukl 1974; Esser and Komorita 1975). We first test this expectation by creating two additional treatment conditions, using a context where there is no guidance available on the accounting issue, but where the audit committee is effective. We choose this context because negotiation strategies likely matter here (given no authoritative guidance and presence of bargaining power by the auditor).

Both new treatments are variations of the *No Guidance/Effective Audit Committee* condition in the main experiment (henceforth referred to as the $C_{no\ AG}$ condition), where the client relents somewhat during the negotiations by being willing to record up to 1 cent per share (i.e., offering some concession). The first additional treatment is the $NC_{no\ AG}$ condition, which varies the initial position of the client (i.e., willing to record up to 1 cent per share) while holding constant its final position during negotiations (i.e., still willing to record only up to 1 cent per share). In the $NC_{no\ AG}$ condition, therefore, the client is essentially not backing off from its initial position (i.e., no concession is offered during negotiations). The second additional treatment condition is the $C_{fn/no\ AG}$ condition, which holds constant the client's initial position (i.e., not willing to record any amount), but varies its final position (i.e., willing to make relevant footnote disclosures on the method used along with the financial impact of using the alternative method proposed by the auditor, without recording any amount of the audit adjustment), thus offering some concession. Our predictions follow from prior research, which shows that "hard" or extreme initial offers followed by relatively small concessions result in better negotiation outcomes for the negotiator

(Chertkoff and Conley 1967; Yukl 1974), and that extreme offers held onto without concessions are least effective (Benton et al. 1972). Hence, we expect that $P_{concede}$ will be lower in the $NC_{no\ AG}$ condition than in the $C_{no\ AG}$ or $C_{fn/no\ AG}$ conditions. However, we do not have a basis to predict any difference in auditors' responses between the $C_{no\ AG}$ and $C_{fn/no\ AG}$ conditions because they differ only in the form of the concession.

Panel A of Table 3 presents the descriptive statistics. Tests of differences between treatment groups indicate no significant differences in auditors' initial recommendations ($p > 0.201$). However, in terms of auditors' perceived negotiation outcome, the proportion of auditors who believe that the amount ultimately booked will result in the client's failure

TABLE 3

Effects of Alternative Treatments on Auditors' Judgments on the Audit Adjustment

Panel A: Descriptive Statistics (in the Context Where There Is No Available Authoritative Guidance, but There is an Effective Audit Committee)

				Initial Recommended Adjustment		Perceived Final Adjustment after Negotiations	
Treatment	**Client's Initial Position**	**Client's Final Position after Negotiations**	**n**	**Proportion[a]**	**Mean (s.d.)[b]**	**Proportion[c]**	**Mean (s.d.)[d]**
$C_{no\ AG}$	Not willing to record any amount	Willing to record up to 1 cent	12	0.58	188 (145)	0.25	117 (127)
$NC_{no\ AG}$	Willing to record up to 1 cent	Willing to record up to 1 cent	13	0.62	192 (144)	0.46	146 (151)
$C_{fn/no\ AG}$	Not willing to record any amount	Willing to make footnote disclosures	11	0.82	245 (121)	0.18	55 (121)

Panel B: Descriptive Statistics (in the Context Where There Is Authoritative Guidance and the Client's Audit Committee is Effective)

				Initial Recommended Adjustment		Perceived Final Adjustment after Negotiations	
Treatment	**Client's Initial Position**	**Client's Final Position after Negotiations**	**n**	**Proportion[a]**	**Mean (s.d.)[b]**	**Proportion[c]**	**Mean (s.d.)[d]**
C_{AG}	Not willing to record any amount	Willing to record up to 1 cent	13	0.85	254 (112)	0.38	154 (127)
NC_{AG}	Not willing to record any amount	Not willing to record any amount	15	0.80	253 (106)	0.47	143 (147)

---

[a] Refers to the proportion of auditors who recommend booking an amount of the audit adjustment that will result in the client's failure to meet the analysts' consensus forecast.

[b] Refers to the mean (standard deviation of) audit adjustment (in $000) that auditors recommend to be booked.

[c] Refers to the proportion of auditors who believe that the amount of the audit adjustment ultimately recorded will result in the client's failure to meet the analysts' consensus forecast.

[d] Refers to the mean (standard deviation of) audit adjustment (in $000) that auditors believe will ultimately be recorded in the client's audited financial statements.

to meet the analysts' consensus forecast is marginally higher in the $NC_{no\ AG}$ condition than in the $C_{no\ AG}$ and $C_{fn/no\ AG}$ conditions combined (proportions = 0.46 and 0.22, respectively; $p = 0.065$, one-tailed). Also, as predicted, $P_{concede}$ is significantly lower in the $NC_{no\ AG}$ condition than in the combined $C_{fn/no\ AG}$ and $C_{no\ AG}$ conditions (proportions = 0.25 and 0.69, respectively; $p = 0.013$, one-tailed).[14] We find no statistically significant difference between the $C_{no\ AG}$ and $C_{fn/no\ AG}$ conditions in terms of auditors' perceived negotiation outcome (proportions = 0.25 and 0.18, respectively; $p = 0.690$) or $P_{concede}$ (proportions = 0.57 and 0.78, respectively; $p = 0.375$). These results support our expectation that auditors have a higher propensity to concede to the client's position in the presence of concessionary moves by the client.

We also examine a third additional treatment, NC_{AG}, where authoritative guidance is available and the audit committee is effective. This treatment is identical to the *Guidance/ Audit Committee Effective* condition in the main experiment (henceforth, the C_{AG} condition) in all aspects except for one major feature: the client does not provide any concessionary move at all during negotiation. As summarized in Panel B of Table 3, in the NC_{AG} condition, the client is not willing to record any amount of the audit adjustment at both stages of the negotiation (i.e., no concession is offered); in contrast, in the C_{AG} condition, the client is unwilling to record any amount at first, but relents to record up to 1 cent per share later (i.e., offering some concession during negotiations). As in the above discussion, we expect $P_{concede}$ to be higher in the C_{AG} condition than in the NC_{AG} condition.[15]

Tests of differences between the C_{AG} and NC_{AG} conditions indicate no significant differences in either the initial recommendations (proportions = 0.85 and 0.80, respectively; $p = 0.748$), perceived negotiation outcome (proportions = 0.38 and 0.47, respectively; $p = 0.330$, one-tailed), or $P_{concede}$ (proportions = 0.55 and 0.42, respectively; $p = 0.267$, one-tailed). These results indicate no effects of concession in the presence of authoritative guidance, and are in contrast to our earlier findings of an effect of concession when authoritative guidance is absent. As an overall test to determine whether the presence of client's concessions interacts with authoritative guidance availability, we perform a categorical analysis with *Client Concessions* (yes, involving treatments $C_{no\ AG}$, C_{AG}, and $C_{fn/no\ AG}$; no, involving treatments $NC_{no\ AG}$ and NC_{AG}) and *Guidance Availability* (yes, involving treatments C_{AG} and NC_{AG}; no, involving treatments $C_{no\ AG}$, $NC_{no\ AG}$, and $C_{fn/no\ AG}$) as independent variables, and either auditors' perceived negotiation outcome or $P_{concede}$ as the dependent variable. Results indicate that the main effect of client concessions is marginally significant for auditors' perceived negotiation outcome ($p = 0.094$, one-tailed) and statistically significant for $P_{concede}$ ($p = 0.022$, one-tailed), but neither the main effect of guidance availability nor the interaction effect is statistically significant ($p > 0.274$).[16] Taken together, the additional analyses presented here provide some evidence that the client's negotiation strategies could exert an influence on auditors' propensity to concede to the client's position. The results do not provide strong evidence that this influence is moderated by the availability of authoritative guidance.

---

[14] Further analyses indicate that the proportion of auditors who believe that the amount ultimately booked will result in the client's failure to meet the analysts' consensus forecast in the $NC_{no\ AG}$ condition is marginally higher than that in the $C_{fn/no\ AG}$ condition (proportions = 0.46 and 0.18, respectively; $p = 0.061$) but not significantly higher than that in the $C_{no\ AG}$ condition (proportions = 0.46 and 0.25, respectively; $p = 0.128$). In contrast, $P_{concede}$ in the $NC_{no\ AG}$ condition is significantly lower than that in the $C_{fn/no\ AG}$ condition (proportions = 0.25 and 0.78, respectively; $p = 0.005$), and marginally lower than that in the $C_{no\ AG}$ condition (proportions = 0.25 and 0.57, respectively; $p = 0.092$, all p-values are one-tailed).

[15] We ask auditors to rate the client's persistence in insisting on their initial position. Persistence ratings are higher in the NC_{AG} condition than in the C_{AG} condition (means = 9.33 and 6.85, respectively; $p = 0.002$, one-tailed). Ratings in the $NC_{no\ AG}$ condition are marginally higher than those in the $C_{no\ AG}$ condition (means = 8.69 and 7.83, respectively; $p = 0.062$, one-tailed), but ratings do not differ between the $C_{fn/no\ AG}$ and $C_{no\ AG}$ conditions (means = 8.09 and 7.83, respectively; $p = 0.715$).

V. DISCUSSION AND CONCLUSION

Using an auditor-client negotiation context, we conduct an experiment involving a quantitatively immaterial audit adjustment that affects the client's ability to meet the analysts' consensus forecast. We find that the availability of authoritative guidance and the effectiveness of the client's audit committee jointly influence auditors' judgments of whether the amount of the audit adjustment ultimately recorded after negotiations with the client will result in the client's failure to meet the analysts' consensus forecast. Our finding suggests that current FASB endeavors that consider the nature of promulgated accounting standards (principles-based versus rules-based guidance) can significantly impact auditor-client negotiations of accounting issues, but that these effects on financial reporting quality are contingent on the effectiveness of the client's audit committee.

Additionally, the results indicate that the availability of authoritative guidance and the effectiveness of the audit committee interact to jointly affect auditors' propensity to concede to the client's preferred position. Results from additional treatment conditions suggest that auditors' propensity to concede to the client's position may be attributed, in part, to the presence of concessionary moves by the client. To the extent that auditors are more likely to objectively assess the merits of the issue before entering into the negotiation process, a requirement to document auditors' (presumably objective) assessments of the issue before the negotiation process, along with close scrutiny and justification of any significant shifts in position from their initial assessments subsequent to negotiations, can serve as a useful control mechanism to mitigate any negative impact of strategic influences that could undermine auditors' effectiveness. Further, the finding that an effective audit committee moderates the impact of authoritative guidance availability (and vice versa) on auditors' perceived negotiation outcome has potential implications for practice in terms of allocation of effort by firms and regulators to compensating mechanisms that serve to jointly promote financial reporting quality.

A limitation of the paper is that it uses a setting where auditors are given general but not specific information about the revenue-recognition issue and the guidance available. Although this allows us to control for context and knowledge differences, it is unclear if the findings generalize to richer settings, an area for future investigation. A good illustration of how this can be successfully done is the study by Hackenbrack and Nelson (1996), which finds that, at least in the context of vague accounting standards, there is convergence in the findings obtained in both sterile and rich settings. One way of extending the current study to a richer context is to use two different revenue-recognition issues that either have or do not have authoritative guidance (because the guidance available for a specific revenue-recognition problem is likely fixed at a given point in time). A trade-off of this approach is that while comparisons can be made on auditors' responses within a specific revenue-recognition issue, caution would have to be exercised with respect to comparisons between revenue-recognition issues, unless differences in ambiguity and complexity of issues are controlled for.

---

[16] As a further analysis, we also conduct a mixed categorical analysis with *Stage* (initial recommendations, perceived negotiation outcome) as a within-subject variable, and *Client Concessions* and *Guidance Availability* as between-subject variables. The results indicate that only the main effect of *Stage* ($p = 0.000$) and the interaction between *Stage* and *Client Concessions* ($p = 0.027$, one-tailed) are statistically significant, while the three-way interaction between *Stage*, *Guidance Availability*, and *Client Concessions* is not statistically significant ($p = 0.402$).

Another limitation is that our participants are audit managers. As we have indicated (footnote 4), audit managers are routinely involved in aspects of client negotiations. They should be aware of what gets documented prior to negotiations, along with negotiated outcomes (see Nelson et al. 2002), which are elements we measure in this study. Still, audit partners are more intimately involved with the client negotiation process, and their responses may not mirror those of managers. In addition, our participants are from one Big 4 firm, and results may not extend to auditors from other audit firms. Also, the study is administered online via the Internet to unobserved participants who perform the experimental task in their normal work environments. As such, it lacks the control typically found in laboratory experiments. Finally, this study examines a situation where recording the full amount of a quantitatively immaterial audit adjustment will result in the client's failure to meet the analysts' consensus forecast by 1 cent per share. Although this provides a boundary condition for testing the effects of authoritative guidance availability and audit committee effectiveness on auditors' judgments, the results may not extend to situations where the analysts' consensus forecast is missed by greater than 1 cent per share.

In addition to the factors examined in this study, future research can further examine the effects of other contextual features (e.g., negotiation expertise) on auditors' judgments and negotiation outcomes. Future research can also delineate the specific mechanisms that influence the strategies (or lack thereof) that auditors employ in their negotiations with clients. Our present design precludes us from doing so. Like Libby and Kinney (2000), we capture auditors' perceived negotiation outcome to provide evidence on how it is jointly influenced by contextual features of auditor-client negotiations. We also measure auditors' initial recommendations, and illustrate the use of a method that provides evidence (in our additional analyses) on how these contextual factors interact to influence auditors' propensity to concede to the client's position. A limitation of our approach is that, while there is greater experimental control, there is no actual interactive negotiation, and the method cannot speak directly on the process that goes on during negotiation. Thus, another useful avenue to extend this line of research is to incorporate actual, interactive dyadic auditor-client negotiations. Although contextually richer, this approach involves trade-offs in terms of the potential noise in the dynamic (and unpredictable) nature of negotiations across dyads, and the cost of assembling participants comprising clients and auditors.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 1999a. *Audit Adjustments.* Statement of Auditing Standards No. 89. New York, NY: AICPA.

———. 1999b. *Audit Committee Communications.* Statement of Auditing Standards No. 90. New York, NY: AICPA.

Anonymous. 1999. AICPA publishes revenue recognition guidance. *The CPA Letter* (Feb/March): 1.

Anonymous. 2002. The FASB address standard overload through new projects. *FASB Report* (February 28, 2002), Financial Accounting Standards Board.

Antle, R., and B. Nalebuff. 1991. Conservatism and auditor-client negotiations. *Journal of Accounting Research* 29 (Supplement): 31–54.

Beasley, M. S., J. V. Carcello, D. R. Hermanson, and P. D. Lapides. 2000. Fraudulent financial reporting: Consideration of industry traits and corporate governance mechanisms. *Accounting Horizons* 14 (December): 441–454.

Benton, A. A., H. H. Kelley, and B. Liebling. 1972. Effects of extremity of offers and concession rate on the outcomes of bargaining. *Journal of Personality and Social Psychology* 24: 73–83.

Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees (BRC). 1999. *Report and Recommendations of the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees.* New Year, NY: NYSE and NASD.

Braun, K. W. 2001. The disposition of audit-detected misstatements: An examination of risk and reward factors and aggregation effects. *Contemporary Accounting Research* 18 (Spring): 71–99.

Brown, C., and I. Solomon. 1990. Auditor configural information processing in control risk assessment. *Auditing: A Journal of Practice & Theory* 10 (Fall): 17–38.

Brown, L. D. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39 (September): 221–241.

Carcello, J. V., and T. L. Neal. 2000. Audit committee composition and auditor reporting. *The Accounting Review* 75 (October): 453–467.

Chertkoff, J. M., and M. Conley. 1967. Opening offer and frequency of concessions as bargaining strategies. *Journal of Personality and Social Psychology* 7: 181–185.

Cohen, J. R., and D. M. Hanno. 2000. Auditors' consideration of corporate governance and management control philosophy in preplanning and planning judgments. *Auditing: A Journal of Practice & Theory* 19 (Fall): 133–146.

Committee of Sponsoring Organizations of the Treadway Commission (COSO). 1999. *Fraudulent Financial Reporting: 1987–1997.* New York, NY: COSO.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (Spring): 1–36.

DeFond, M. L., and J. Jiambalvo. 1991. Incidence and circumstances of accounting errors. *The Accounting Review* 66 (July): 643–655.

Esser, J. K., and S. S. Komorita. 1975. Reciprocity and concession making in bargaining. *Journal of Personality and Social Psychology* 31: 864–872.

Fox, J. 1997. Learn to play the earnings game (and Wall Street will love you). *Fortune* (March 31): 76–80.

Gibbins, M., S. Salterio, and A. Webb. 2001. Evidence about auditor-client negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (December): 535–563.

———, and ———. 2002. Modeling the auditor's intended strategy in auditor-client negotiation. Working paper, University of Alberta and University of Waterloo.

Hackenbrack, K., and M. Nelson. 1996. Auditors' incentives and their application of financial accounting standards. *The Accounting Review* 71 (January): 43–59.

Johnstone, K. M., M. H. Sutton, and T. D. Warfield. 2001. Antecedents and consequences of independence risk: Framework for analysis. *Accounting Horizons* 15 (March): 1–18.

Kennedy, J., D. N. Kleinmuntz, and M. E. Peecher. 1997. Determinants of the justifiability of performance in ill-structured tasks. *Journal of Accounting Research* 35 (Supplement): 105–123.

Knapp, M. C. 1987. An empirical study of audit committee support for auditors involved in technical disputes with client management. *The Accounting Review* (July): 578–588.

Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, New York, September 28.

Libby, R., and W. Kinney. 2000. Does mandated audit communications reduce opportunistic corrections to manage earnings to forecasts? *The Accounting Review* 75 (October): 383–404.

———, R. Bloomfield, and M. W. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27 (November): 775–810.

McAlister, L., M. H. Bazerman, and P. Fader. 1986. Power and goal setting in channel negotiations. *Journal of Marketing Research 23* (August): 228–236.

McMullen, D. A. 1996. Audit committee performance: An investigation of the consequences associated with audit committees. *Auditing: A Journal of Practice & Theory* 15 (Spring): 87–103.

National Association of Corporate Directors (NACD). 1999. *Report of the NACD Blue Ribbon Commission on Audit Committees*. Washington, D. C.: NACD.

Neale, M. A., and G. B. Northcraft. 1991. Behavioral negotiation theory: A framework for conceptualizing dyadic bargaining. *Research in Organization Behavior* 13: 147–190.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings-management decisions. *The Accounting Review* (Supplement): 175–202.

Nichols, D. R., and K. H. Price. 1976. The auditor-firm conflict: An analysis using concepts of exchange theory. *The Accounting Review* 51 (April): 335–346.

Pruitt, D. G., and P. J. Carnevale. 1993. *Negotiation in Social Conflict.* Pacific Grove, CA: Brooks/ Cole.

Salterio, S, and L. Koonce. 1997. The persuasiveness of audit evidence: The case of accounting policy decisions. *Accounting, Organizations and Society* 22: 573–587.

Securities Exchange Commission (SEC). 1999a. *Revenue Recognition in Financial Statements.* SEC Staff Accounting Bulletin: No. 101. December 3. Washington, D.C.: SEC.

———. 1999b. *Materiality*. SEC Staff Accounting Bulletin: No. 99. August 12. Washington, D.C.: SEC.

Siegel, S., and N. Castellan. 1988. *Nonparametric Statistics for the Behavioral Sciences*. New York, NY: McGraw-Hill.

Trompeter, G. 1994. The effect of partner compensation schemes and generally accepted accounting principles on audit partner judgment. *Auditing: A Journal of Practice & Theory* 13 (Fall): 56–68.

Yukl, G. A. 1974. Effects of situational variables and opponent concessions on a bargainer's perceptions, aspirations, and concessions. *Journal of Personality and Social Psychology* 29: 227–236.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 819–846

Materiality Uncertainty and Earnings Misstatement

Evelyn R. Patterson
State University of New York at Buffalo

Reed Smith
Indiana University

ABSTRACT: The concept of materiality provides a basis for auditors to ignore small misstatements, but the definition of "small" in this context is ambiguous. The issue of "materiality-standard-setting" has been raised recently by Arthur Levitt, former chairman of the Securities and Exchange Commission. We examine how materiality uncertainty affects the auditor's evaluation of audit evidence and a manager's choice of earnings overstatement in a strategic auditing model where earnings misstatements also include unintentional system error. We find that when the expected cost of accepting financial statements that are materially misstated, which we refer to as an audit failure, is relatively large, an increase in materiality uncertainty results in a more conservative auditor evaluation of the audit evidence and a decrease in the amount of intentional overstatement. Alternatively, if the auditor's expected cost of extending audit procedures is relatively high, then an increase in materiality uncertainty results in a less conservative auditor evaluation of the audit evidence and an increase in the manager's earnings overstatement. The auditor also becomes increasingly conservative as the report increases when the information system is sufficiently noisy. Finally, when the expected cost of audit failure is large, the equilibrium audit risk can increase *or* decrease in materiality uncertainty despite the corresponding increase in auditor conservatism and decrease in intentional overstatement. Audit risk is the average probability of audit failure across all possible evidence outcomes.

Keywords: *strategic audit; materiality; intentional overstatement; system of internal control.*

I. INTRODUCTION

The issue of materiality has recently garnered renewed attention from the Securities and Exchange Commission (SEC) and the accounting profession. In August 1999 the SEC issued Staff Accounting Bulletin No. 99 (SAB No. 99), *Materiality*. The purpose of this bulletin is not to adopt precise materiality standards, but to provide "guidance in applying materiality thresholds to the preparation of financial statements filed with

---

The authors appreciate the helpful comments of Paul Newman, Rajib Doogar, Bill Heninger, Bill Kinney, Dale Morse, Kenny Reynolds, Mike Stein, and participants at the American Accounting Association 2001 Midyear Auditing Conference. We give special thanks to Richard Sansing and to two anonymous reviewers for their helpful suggestions.

Editor's note: This paper was accepted by Richard C. Sansing, Special Editor.

Submitted July 2001
Accepted February 2003

the Commission and the performance of audits of those financial statements." Prior to 1999, auditors primarily referred to Statement of Financial Accounting Concepts (SFAC) No. 2, which defines a material misstatement as one such that "the judgment of a reasonable person relying on the information would have been changed or influenced by the...misstatement" (paragraph 10).[1] The SEC has been concerned about public companies "abusing materiality guidelines in order to manipulate their reported earnings" (*Journal of Accountancy* 2000, 17–18)—that is, the influence of materiality on intentional earnings misstatement.

We examine, in a stylized, game-theoretic model, how uncertainty regarding the materiality threshold affects the interaction between an auditor and manager when the manager has the opportunity and incentive to intentionally overstate earnings. The manager chooses an amount of intentional overstatement. The auditor evaluates the audit evidence by accepting or rejecting reported earnings as materially correct.[2] In addition to intentional overstatements, reported earnings misstatements contain unintentional system errors. The manager has no private information about earnings and overstates earnings by biasing the information system report on earnings. The auditor observes reported earnings before evaluating the audit evidence. The report allows the auditor to update his expectation of materiality and unintentional system error.

While the SEC has identified materiality as a major concern, it is reluctant to provide precise guidance on the issue. Former SEC Chairman Arthur Levitt states that "materiality is not a bright-line cutoff of 3% or 5%. It requires consideration of all relevant factors that could impact an investor's decision." (See *WSJ* 1998.)[3] Neither the professional literature, common law, nor statutory law provides a precise materiality standard. However, the impact of a vague definition of materiality is unknown, particularly when intentional misstatements are possible. Previous studies in statistical auditing (see for example, Johnstone 1995), decision theory (see for example, Kinney 1975) and strategic auditing (see for example, Newman and Noel 1989) have assumed that a precise materiality standard exists.

Uncertainty about the materiality threshold (materiality uncertainty) arises from the judgmental nature of materiality and potentially creates additional tensions in the strategic interaction between the auditor and manager. While both the auditor and manager have expectations about the materiality threshold, the *actual* threshold is affected by qualitative factors such as the quality of accounting estimates.[4]

We find that the effect of materiality uncertainty on the equilibrium strategies depends on the auditor's expected cost of accepting materially misstated financial statements, which we refer to as an audit failure, relative to the expected cost of rejection. The expected cost of rejection includes the cost of extending audit procedures and the resulting loss of client goodwill or lucrative consulting contracts. Materiality uncertainty affects the auditor's evidence evaluation because it affects the updated probability of material misstatement (for a fixed earnings report and audit evidence). When the cost of audit failure is high relative to the cost of rejection, the auditor evaluates the evidence more conservatively because he has a greater incentive to avoid audit failure.[5] Greater uncertainty about the materiality threshold increases the updated probability of material misstatement, exposing the auditor to a greater risk of audit failure. Consequently, the auditor evaluates the evidence more conservatively for higher levels of materiality uncertainty, which, in turn, induces the manager to decrease the amount of intentional overstatement. The opposite effects occur when the cost of audit failure is relatively lower than the cost of rejection. In this case, the manager increases the amount of intentional overstatement and the auditor evaluates the evidence less conservatively.

---

[1] This is the definition adopted by the Auditing Standards Board in Statement on Auditing Standards (SAS) No. 47, *Audit Risk and Materiality in Conducting an Audit* (AICPA 1984).

[2] As we discuss in Section II, rejection in our setting implies that the auditor extends audit procedures rather than actually issuing a modified report.

[3] Proponents of a precise materiality standard argue that the lack of a precise definition induces non-uniformity in the presentation of financial information and makes it impossible for companies to know what is expected of them. Jennings et al. (1985) argue that "[t]he accounting profession should be strongly encouraged to revive their efforts to establish explicit standards in the materiality area, and the legal profession should actively contribute to that process." In addition, ambiguity not only creates a problem for preparers and users, but also for enforcement and for the courts. John Fedders (1998), the former Director of Enforcement for the SEC, asserts that "[t]he tendency of the law must always be to narrow the field of uncertainty, so that citizens may know by what standards their conduct will be judged." He argues that materiality has been impossible to implement from an enforcement perspective because of its ambiguity.

[4] See Section II for a more detailed discussion of the qualitative aspects of materiality.

Anecdotal evidence suggests that the auditor's expected cost of audit failure is typically large relative to the expected cost of extending procedures. Historically, the audit profession has focused on developing professional standards that promote the avoidance of audit failure. For example, the independence rules in the Code of Professional Ethics prohibit auditors from holding direct investments (even one share of stock) in their client firms (AICPA 2001). This reduces the auditor's incentive to evaluate the evidence less conservatively. In addition, damage awards against auditors for undetected material misstatements have been large. For example, Coopers & Lybrand settled its case with debenture holders of Miniscribe for $140 million and claims against Arthur Andersen for Enron are estimated to be in the billions of dollars.[6] Furthermore, legislative action has addressed the possibility of large expected rejection costs due to lucrative consulting contracts. The Sarbanes-Oxley Act, signed into law on July 30, 2002, prohibits corporations from hiring the same accounting firms for audits and most types of consulting services.[7] To the extent that the expected cost of audit failure is large relative to the expected cost of extending audit procedures, our model suggests that a fixed materiality standard would be detrimental in promoting the avoidance of audit failure.

Though increasing materiality uncertainty induces more conservative evidence evaluation and less intentional overstatement by the manager when the expected cost of audit failure is large, equilibrium audit risk can increase *or* decrease. Audit risk is the *average* probability of audit failure across all possible evidence outcomes. Increased auditor conservatism has an indirect negative effect on audit risk. However, this indirect effect may be overwhelmed by the direct positive effect of materiality uncertainty on audit risk for lower levels of materiality uncertainty. For lower levels of materiality uncertainty, the indirect marginal effect of a change in materiality uncertainty on audit risk is relatively small.

We also find that if the information system is sufficiently noisy, the auditor's evidence evaluation strategy becomes increasingly more conservative as reported earnings increases. This result holds despite the fact that higher earnings reports also imply that the expected materiality threshold has increased. If the system is sufficiently noisy, then increases in the report imply that the updated probability of an audit failure also increases, so the auditor evaluates the evidence more conservatively.

Our model most resembles that of Newman and Noel (1989) because we consider two strategic choices very similar to theirs. The auditor chooses an evidence evaluation strategy that is based upon sample information, and the manager chooses an amount of intentional overstatement that affects the distribution of audit evidence. There are two key differences between our model and Newman and Noel (1989). First, the manager in our setting is not limited to a binary choice of "no fraud" and "fraud" that the auditor finds "acceptable" and "unacceptable," respectively. As a result, an uncertain materiality threshold is not possible in their model. Second, we introduce an additional source of reported earnings misstatement—unintentional system error. Neither the auditor nor the manager knows the amount of misstatement due to unintentional system error. Further, the auditor in our model is concerned about total earnings misstatement consisting of intentional and unintentional error.

---

[5] When the auditor evaluates the evidence more conservatively, he rejects earnings as materially correct for evidence observations that previously would have resulted in acceptance.

[6] See Kiernan and Lewin (1994) and Trager (2002).

[7] This is also referred to as the Anti-Fraud bill. See *Engineering News-Record* (2002).

In Section II, we describe the players' strategies as well as the economic consequences of those strategies. In Section III we discuss our solution technique and determine the equilibrium strategies, while Section IV provides a comparative analysis of equilibrium strategies and audit risk. We present concluding remarks in Section V.

II. MODEL DESCRIPTION

When formulating an audit plan, the auditor considers the motivation and opportunity of the manager to commit fraud and whether potential misstatements materially affect financial statement fairness. After observing an earnings report and obtaining audit evidence, the auditor either accepts the earnings report as materially correct or rejects, extending audit procedures. As in previous strategic audit settings, the manager in our setting obtains benefits from intentional misstatements that are not identified by audit procedures. Consistent with this literature, we also assume that larger intentional misstatements increase both the likelihood that the misstatements are identified and the manager's expected penalty for misstatements if the auditor rejects. The auditor, on the other hand, must balance the expected cost of audit failure against the expected cost of extending audit procedures.

When we add materiality to this interaction, the tensions within the model change because the manager chooses intentional misstatement knowing that the auditor considers some of them too small to affect financial statement fairness. To explore the strategic effects of materiality, we examine a stylized model in which the manager is limited to overstating earnings by biasing an accounting information system report, while the auditor employs an evidence evaluation strategy. We describe specific elements of our stylized model in the following subsections.

The Accounting Information System and Earnings Misstatement

The manager chooses an amount of overstatement bias, denoted ω, in anticipation of an earnings report and an audit, but before either event occurs. The manager knows that the distribution of true earnings is $\tilde{y} \sim N(\mu_y, \sigma_y^2)$. He also knows that the accounting information system, in the absence of any intentional misstatement, would generate an earnings report, z, that is equal to true earnings plus noise, $\tilde{e}$, where $\tilde{e} \sim N(0, \sigma_e^2)$. We assume that earnings and system error are independent random variables. Thus the earnings report is:

$$\tilde{z} = \tilde{y} + \tilde{e} + \omega, \tag{1}$$

where $\tilde{z} \sim N(\mu_y + \omega, \sigma_y^2 + \sigma_e^2)$. We denote the density of $\tilde{z}$ as $\lambda(z)$.

The manager adds intentional overstatement bias, ω, to the earnings report by misreporting selected transactions. This strategy is advantageous to the manager for two reasons. First, he can inflate earnings without obvious large or unusual year-end adjustments.

Second, he can incorporate transaction errors that normal audit procedures do not characterize as intentional or unintentional. For example, suppose two types of errors can occur in the recording of sales: (1) goods are shipped that have not been billed—creating an understatement error or (2) goods are billed that have not been shipped—creating an overstatement error. To bias revenue upward, the manager increases the frequency of the latter error type. The manager could sell goods to some customers as "bill and hold," claiming that the goods are sold but are being held for the customer. He could also send more goods than the customer requested or bill sales based on customer orders, before the shipment of goods occur.[8] Confirmations might uncover these as recording errors, but the manager could claim that they were unintentional system errors.[9]

As a result of the earnings overstatement bias and the system error, the total earnings misstatement included in an earnings report is:

$$\tilde{\theta} = \tilde{z} - \tilde{y} = \tilde{e} + \omega. \tag{2}$$

Because ω is not stochastic, the total misstatement is distributed, $\tilde{\theta} \sim N(\omega,\sigma_e^2)$.

The auditor observes the report z before he performs audit procedures. From the auditor's point of view, the earnings misstatement is:

$$\tilde{\theta}_z = z - \tilde{y}_z = \tilde{e}_z + \omega, \tag{3}$$

where the subscript, z, denotes that the random variable is conditioned on the observed earnings report. Applying Bayes theorem, we find that $\tilde{y}_z \sim N(\mu_{y|z},\sigma_{y|z}^2)$ where $\mu_{y|z} = \delta(z - \omega) + (1 - \delta)\mu_y$, $\delta = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_e^2}$ and $\sigma_{y|z}^2 = \frac{\sigma_y^2\sigma_e^2}{\sigma_y^2 + \sigma_e^2}$.[10] In other words, once the auditor observes the report z, he revises his expectation of earnings by weighting the unbiased system report, $z - \omega$, and the unconditional expectation of earnings, μ_y, using their relative variances. Consequently, the total misstatement conditioned on the report is $\tilde{\theta}_z \sim N(z - \mu_{y|z},\sigma_{y|z}^2)$ where $z - \mu_{y|z} = \omega + (1 - \delta)(z - \omega - \mu_y)$.

Materiality of an Earnings Misstatement

According to AU§312, AICPA (2001), "Audit Risk and Materiality in Conducting an Audit," the auditor should take into account both quantitative and qualitative aspects of misstatements when considering whether they are material. For example, a $300,000 misstatement that is .001 percent of net assets and 5 percent of earnings may be material because users rely on earnings in making investment decisions. Qualitative factors also include items such as the regulatory requirements of the firm. According to SAB No. 99, auditors should not rely entirely on quantitative measures. However, quantitative measures (such as 5 percent of earnings) are important characteristics in judging the materiality of misstatements (see AU§9312.11 [AICPA 2001] for discussion). Moreover, percentage-of-earnings is often used by auditors as a quantitative measure of materiality (see Ricchiute 2002, 52).

---

[8] See Konrath (2001, 719–721) for an example of an information system with this feature. Anecdotally, turnkey software packages often generate a sales invoice (i.e., record a sale) prior to evidence of the shipment of goods. Thus, the system relies on a control that purges a sale if no shipment occurs.

[9] Despite the choice of overstatement bias by the manager prior to the realization of earnings, the manager is likely to have private information about the information system or the earnings distribution. For simplicity, we do not model this aspect of the manager's information.

[10] See Berger (1985, 127–128) for details on deriving the posterior normal distribution.

To incorporate the judgmental characteristics (both quantitative and qualitative) of materiality, we assume that the materiality threshold of reported earnings is a normally distributed random variable $\tilde{m}_z \sim N(M_z,\sigma_m^2)$. We denote this density $t(m_z)$.[11] The subscript, z, indicates that the materiality measure depends on the magnitude of reported earnings, z.[12] To simplify the model, we assume that the earnings report affects only the mean of the materiality distribution. In other words, $M_z = \pi z$ where π is an exogenous percentage parameter. For example, in an earnings report of $20,000 and given $\pi = 5\%$, the auditor's materiality expectation, M_z, equals $1,000. As the earnings report increases, the expected materiality threshold increases.

Our setting includes the possibility of both intentional and unintentional earnings misstatements. Regarding the distinct nature of these two types of misstatements, the professional literature does not suggest that the auditor plan procedures to separately determine material misstatements of each type. Rather AU§312.05, AICPA (2001), points out that "[t]he auditor has *no* responsibility to plan and perform the audit to obtain reasonable assurance that misstatements, whether *caused by errors or fraud*, that are not material to the financial statements are detected" (emphasis added). Furthermore, AU§312.08, AICPA (2001), states, "When considering the auditor's responsibility to obtain reasonable assurance that the financial statements are free of material misstatement, there is no important distinction between errors and fraud."

In our model there is a single distribution over materiality that does not distinguish between (unintentional) errors and fraud (intentional errors). The auditor considers all misstatements, θ_z, whether due to error or fraud, greater than or equal to m_z as material.

Audit Evidence

We assume that the auditor observes a noisy evidence signal:

$$\tilde{x}_z = \tilde{\theta}_z + \tilde{\xi} \tag{4}$$

of the total misstatement $\tilde{\theta}_z$ where $\tilde{\xi} \sim N(0,\sigma_x^2)$ and $\tilde{\theta}_z$ and $\tilde{\xi}$ are independent. Thus, $\tilde{x}_z \sim N(z - \mu_{y|z},\sigma_{y|z}^2 + \sigma_x^2)$. We denote the density of $\tilde{x}_z$ as $f(x_z)$ and the distribution function as $F(x_z)$. Evidence, x_z, is the outcome of an audit procedure such as a detail test of a sample of transactions or analytical review. The evidence cannot distinguish between intentional and unintentional misstatements but as x_z increases, the auditor infers that material misstatement is more likely. Normal, non-forensic-type audit procedures are often limited in their ability to detect a specific source of misstatement. As discussed above, the auditor in our model is concerned about whether, in total, a material misstatement exists, "due to error or fraud."[13]

---

[11] Our assumption that the materiality threshold is normally distributed is made for technical convenience. We have no reason to assume that the uncertainty about the threshold is represented by any specific distribution. The assumption of normality allows us to examine how the variance (or degree of uncertainty) affects the interaction between the auditor and manager in our model. In addition, we recognize that this assumption admits some unrealistic values for the materiality threshold, such as negative amounts. This limitation of our assumption does not affect our results qualitatively.

[12] In determining materiality, the auditor typically uses *reported* values as a base amount (see Ricchiute [2002, 50–58 and 67] for discussion). Of course, materiality based on expected earnings is also reasonable and our results are not affected by considering this measure of materiality.

[13] This is typically how an audit is performed. For example, if the evidence is obtained from analytical review of sales, then the auditor is only concerned with the difference between total predicted sales and total recorded sales. Transaction misstatements that net to zero do not overstate recorded sales nor do they provide the manager

(*continued on next page*)

Once the auditor observes x_z, he updates the probability of misstatement. The random variable for misstatement, conditional on both x_z and z, is $\tilde{\theta}_{x_z,z} \sim N(\mu_{\theta|x_z,z}, \sigma_{\theta|x_z,z}{}^2)$, where $\sigma_{\theta|x_z,z}{}^2 = (1 - \alpha)\sigma_{y|z}{}^2$, $\mu_{\theta|x_z,z} = \alpha x_z + (1 - \alpha)(z - \mu_{y|z})$ and $\alpha = \frac{\sigma_{y|z}{}^2}{\sigma_{y|z}{}^2 + \sigma_x{}^2}$. Based on the updated misstatement the auditor decides whether to accept the earnings report as materially correct or to extend audit procedures. Extended procedures may consist of additional standard audit procedures such as confirmation or forensic-type tests such as authenticating documents.[14]

Finally, we assume that the system is sufficiently noisy relative to the mean materiality percentage π and the precision of the audit evidence $\sigma_x{}^2$. Specifically, we assume $\sigma_\varepsilon{}^2 > \frac{\pi\sigma_x{}^2\sigma_y{}^2}{(1 - \pi)(\sigma_x{}^2 + \sigma_y{}^2)}$, so that our game has an interior solution. While we make this assumption for tractability reasons, the assumption also makes economic sense for several reasons. First, the accuracy of information systems is limited due to cost. Second, the mean materiality percentage π is likely to be small (less than 10 percent, see Ricchiute [2002, 50–55]). Third, the cost of audit evidence is also likely to be sufficiently small ($\sigma_x{}^2$ is sufficiently small) so that this assumption is met.

The Players' Payoffs

The Auditor

There are two tensions guiding the auditor's choice of accepting the earnings report or extending audit procedures. If the auditor accepts an earnings report that is materially overstated, $(\theta_z - m_z \geq 0)$, then he incurs a cost of audit failure equal to L. L incorporates all expected costs of a type II error, including potential litigation, sanctions, and reputation loss. After observing the audit evidence and the earnings report, the auditor updates the probability that the earnings misstatement is material. Let $H(\cdot|x_z)$ denote the cumulative distribution for the earnings misstatement less materiality, $\tilde{\theta}_z - \tilde{m}_z$, conditional on the observed evidence, x_z. The auditor's expected payoff for accepting earnings as materially correct, given the audit evidence x_z and an earnings report z is:

$$-(1 - H(0|x_z))L \tag{5}$$

and $H(\cdot|x_z)$ is the distribution function of the random variable $\tilde{\theta}_{x_z,z} - \tilde{m}_z \sim N(\mu_{\theta|x_z,z} - M_z, \sigma_{\theta|x_z,z}{}^2 + \sigma_m{}^2)$. The term, $(1 - H(0|x_z))$ represents the probability that the financial statements are materially misstated given the report z and the audit evidence, x_z. That is, $(1 - H(0|x_z))$ is the updated, *ex post* measure of audit risk given the auditor accepts.[15]

If the auditor chooses to reject the report in favor of extending audit procedures, then he incurs a cost, R, that reflects the expected costs of the additional work and the potential

---

Footnote 13; continued
with incremental benefits in our model. Our model is not intended to address illegal acts such as illegal payments where individual small amounts can affect the financial statements in a material way. Similarly, our model does not consider the audit of multiple accounts (a multilocation problem) within earnings where the auditor is concerned about the materiality of account classification. Newman et al. (1996), for example, examine a multilocation audit setting but do not include materiality in their analysis. We leave the study of materiality in a multilocation setting for future research.

[14] The Public Oversight Board (POB) recently issued "The Panel on Audit Effectiveness: Report and Recommendations" in 2000 that suggests auditors incorporate forensic-type procedures into the audit plan. If the auditor is not trained to perform forensic-type tests that are warranted, then the auditor should consider engaging a Certified Fraud Examiner.

[15] See footnote 28 for a discussion of *ex ante* and *ex post* audit risk measures.

bad will that is generated with the client.[16] Typically, L is considered to be large compared to R because auditors expect to incur large litigation costs for undiscovered material misstatements.[17] Most auditors are relatively more concerned with avoiding L relative to R.

The auditor extends audit procedures when:

$$L(1 - H(0|x_z)) \geq R. \tag{6}$$

As a result, the updated probability of audit failure never exceeds R/L.[18]

Based on the Neyman-Pearson Lemma, the auditor's evaluation strategy can be represented as a cut-off value c_z such that $x_z < c_z$ results in acceptance and $x_z \geq c_z$ results in extending procedures.[19] Consequently, the auditor's expected payoff based on his choice of c_z for a given z is:

$$A = -L \int_{-\infty}^{c_z} (1 - H(0|x_z))f(x_z)\, dx_z - R(1 - F(c_z)) \tag{7}$$

where $-L \int_{-\infty}^{c_z} (1 - H(0|x_z))f(x_z)\, dx_z$ is his expected payoff from acceptance and $-R(1 - F(c_z))$ is his expected payoff from extending procedures.

The Manager

The manager's expected payoff depends on reported earnings, which includes both intentional and unintentional misstatements. Because the manager chooses overstatement bias, ω, before the earnings report, z, is realized, he calculates his expected payoff by anticipating how the earnings report affects his payoff for each report z and by anticipating how the auditor will evaluate the evidence for each observed x_z.

The manager receives a payoff proportional to the earnings report z but incurs an expected penalty of $p\omega$ if the auditor extends procedures. The penalty multiplier, p, captures the relative weightings on the positive impact of the earnings report and the negative impact of the penalty.[20] As a result, the manager anticipates an expected penalty that is proportional to ω whether the amount of intentional misstatement is material or immaterial. This is consistent with SAS No. 82 (AICPA 1997, para. 38) that says evidence of "fraud involving senior management...should be reported directly to the audit committee." Finally, we assume that the manager does not benefit from corrected misstatements arising from auditor rejection.[21] Once the auditor rejects, the manager's compensation (based on z) is reduced by $p\omega$, which represents fines, sanctions, or incarceration for criminal behavior.[22]

---

[16] We assume that all identified misstatements are corrected. This is quite common when audit procedures are extended. Extending audit procedures is a form of "announcement" that audit problems likely exist. Moreover, because the client incurs costs for the extra work, the client is usually compelled to make corrections for misstatements the auditor finds. Of course, the remote possibility exists that the auditor will find misstatements so negligible that no adjustments are made. We do not model the technical details of this stage of the game and leave it for future research.

[17] See Kiernan and Lewin (1994) and Trager (2002).

[18] Note that if $R > L$, then extending audit procedures is a dominated strategy and the auditor always accepts. Because we do not observe this behavior in practice, R is likely to be less than L.

[19] See Berger (1985, 523–530). We technically solve our game as a simultaneous move game in which the auditor chooses c_z and the manager chooses ω. We could alternatively state that the auditor chooses c_z at the beginning of the game, observes z, and then compares the stochastic signal to c_z in deciding whether to accept or reject.

[20] For example, the manager may receive a 10 percent bonus based on reported earnings and be penalized twice the overstatement if caught. The net effect of these two multiples is captured in p. The value of p also captures the possibility that the penalty may not be levied with certainty.

The manager chooses system bias prior to the realization of z and x_z. Consequently, he averages his expected payoff over all possible reports. For each earnings report z, the manager considers the probability that $x_z \geq c_z$ in assessing the likelihood that audit procedures are extended.[23] The manager's expected payoff is:

$$U = \int_{-\infty}^{\infty} \{z - p\omega(1 - F(c_z))\}\lambda(z)\, dz. \tag{8}$$

We assume that if the manager chooses $\omega < 0$, then he is either penalized zero or some positive amount. As a result, the manager chooses a non-negative system bias that potentially overstates reported earnings.

III. EQUILIBRIUM STRATEGIES

The solution concept for our model is a Bayesian-Nash equilibrium (see Fudenberg and Tirole 1995). Each player chooses his Nash strategy, given his beliefs about the earnings report and the equilibrium choice of the other player. We solve for the players' equilibrium strategies by first considering the auditor's strategy as a function of the earnings report z and the observed evidence x_z, fixing the manager's overstatement bias, ω.

Both x_z and z provide the auditor with information about the updated expected earnings misstatement. The auditor maximizes his expected payoff by extending procedures when the expected cost of acceptance exceeds the cost of extending audit procedures as shown in Expression (6).

Expression (6) implies that the auditor extends audit procedures whenever:

$$\mu_{\theta|x_z,z} \geq M_z - \Phi^{-1}\left(\frac{L - R}{L}\right)\sqrt{(1 - \alpha)\sigma_{y|z}^{\ 2} + \sigma_m^{\ 2}} \tag{9}$$

where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal distribution. The auditor compares the updated expected earnings misstatement (likely misstatement) to expected materiality less a constant that depends on the variance of the earnings misstatement, the variance of materiality and the auditor's relative costs for type I and type II errors. If $L = 2R$, then the auditor decides to extend procedures if and only if likely misstatement exceeds expected materiality. However, the expected cost of audit failure is typically large ($L > 2R$). In this case, the auditor extends audit procedures if likely misstatement exceeds an amount that is *less* than the expected materiality threshold. The effect of the manager's strategy and expected materiality on the likelihood of extending procedures is intuitive. As the amount of manager bias increases (which increases the mean of total misstatement $\mu_{\theta|x_z,z}$) or expected materiality decreases, the auditor is more likely to extend procedures.

---

[21] Our model does not include the potential negotiation between the auditor and manager involved in determining whether identified misstatements result in required adjustments by the auditor. Our setting focuses on the identification of fraud when it exists and procedures are extended. Nevertheless, our model captures the concerns of Arthur Levitt and the practitioner literature based on the manager's *anticipation* of the size of *likely* misstatement that the auditor finds acceptable. In discussing the evaluation of audit findings, AU§312.27, AICPA (2001) notes that "[I]n evaluating whether the financial statements are presented fairly, in all material respects...the auditor should aggregate misstatements that the entity has not corrected...that enables him to consider whether...they materially misstate the financial statements." AU§312.28, AICPA (2001) goes on to say that "[t]he aggregation of misstatements should include the auditor's best estimate of the total misstatements...(hereafter referred to as likely misstatement), not just the amount of misstatements he specifically identifies." Our model captures this portion of the evaluation process in which the auditor projects likely misstatement and must decide the likelihood that the misstatement is material to reported earnings.

[22] For simplicity, we assume that all managers are prone to fraud. As a result, the auditor can compute the amount of fraud in equilibrium. This assumption is made without loss of generality though. As discussed in Newman et al. (1996), even though the auditor may be able to infer the amount of fraud no manager penalties can be assessed unless the auditor obtains evidence of fraud. Furthermore, our results remain qualitatively unchanged if we add an inherently honest manager to the game.

[23] Clearly, p must be greater than 1. Otherwise, the manager has no incentive to limit overstatement bias.

When the auditor observes both x_z and z, his decision criterion can be expressed as a relation between x_z and a function of z denoted c_z where the auditor accepts if $x_z < c_z$ and rejects if $x_z \geq c_z$. We find the evaluation strategy c_z by solving for the x_z that results in auditor indifference between accepting and extending procedures. This c_z maximizes the auditor's expected payoff in Expression (7) for each report z. Thus in equilibrium, the auditor's strategy for evaluating the audit evidence x_z is equal to:

$$c_z = \frac{\pi z - (1 - \alpha)(z - \mu_{y|z}) - \Phi^{-1}\left(\frac{L - R}{L}\right)\sqrt{(1 - \alpha)\sigma_{y|z}^2 + \sigma_m^2}}{\alpha}. \tag{10}$$

Expression (10) shows how the auditor's updated information about earnings, $\mu_{y|z}$, affects his evaluation strategy. A higher earnings report indicates a higher expected conditional earnings $\mu_{y|z}$. However, a higher earnings report also increases the *expected* misstatement since $(z - \mu_{y|z}) = \omega + (1 - \delta)(z - \omega - \mu_y)$ is increasing in z.[24]

We derive the equilibrium strategies of the auditor and manager by expressing the auditor's strategy in a linear form and then solving for the manager's first order condition. Proposition 1 provides the unique equilibrium strategies of the auditor and manager.

Proposition 1: The unique interior equilibrium strategies of the auditor and manager are characterized as follows.

1) The auditor chooses $c_z = az + b$ with:

$$a = \frac{\pi - (1 - \alpha)(1 - \delta)}{\alpha}$$

and

$$b =$$

$$\frac{-(1 - \alpha)(\delta\omega - (1 - \delta)\mu_y) - \Phi^{-1}\left(\frac{L - R}{L}\right)\sqrt{(1 - \alpha)\sigma_{y|z}^2 + \sigma_m^2}}{\alpha}.$$

2) Given a and b defined in part (1), the manager chooses overstatement bias, ω, that satisfies:

$$1 - p(1 - \mathcal{F}(c)) - (1 - a)p\omega f(c) = 0 \tag{11}$$

---

[24] We examine the equilibrium effect of the earnings report on c_z in Proposition 4.

where $(1 - a) > 0$, $c = a(\omega + \mu_y) + b$ and $\mathcal{f}(\cdot)$ is the density and $\mathcal{F}(\cdot)$ is the distribution function of the unconditional evidence signal, $\tilde{x} \sim N(\omega, (1 - a)^2\, \sigma_e + a^2\sigma_y^2 + \sigma_x^2)$.

All proofs are in the Appendix.

Proposition 1 part (1) shows that the auditor's evaluation strategy is linear in z. For a given z, he responds to the manager's choice of overstatement bias by adjusting c_z proportionate to ω. This is consistent with the fact that the manager chooses overstatement bias before the realization of y and his choice cannot be affected by y.

We obtain the manager's equilibrium condition in Expression (11) by substituting the auditor's evaluation strategy c_z in its general form, $az + b$, into the manager's payoff function, integrating over z and then differentiating with respect to ω. The manager's choice of overstatement bias takes into account the marginal (*ex ante*) effect on his payoff $(1 - p(1 - \mathcal{F}(c)))$ plus the marginal effect of ω on the probability of rejection $(-(1 - a)p\omega \mathcal{f}(c))$.

IV. COMPARATIVE ANALYSIS

In this section, we provide a comparative analysis in order to explore the interactions of the manager and the auditor. Specifically, we consider how materiality uncertainty and the earnings report affect the equilibrium strategies and audit risk.

The Effect of Materiality Uncertainty on the Equilibrium Strategies

Materiality measures how misstatements affect the decisions of users and, as a result, the size of misstatement the auditor is willing to accept as inconsequential to financial statement fairness. Given that materiality is an inherently uncertain quantity, we perform a comparative analysis of how uncertainty in the materiality threshold affects the players' equilibrium strategies.

Proposition 2 shows how the auditor's choice of evaluation strategy and the manager's choice of overstatement bias change with the degree of materiality uncertainty.[25]

Proposition 2: In equilibrium:

1) the auditor's evaluation strategy, c_z, and the manager's choice of overstatement bias, ω, decrease in materiality uncertainty, σ_m, if and only if $2R < L$.
2) the auditor's evaluation strategy, c_z, and the manager's choice of overstatement bias, ω, increase in materiality uncertainty, σ_m, if and only if $2R > L$.

Some proponents of a well-defined materiality standard argue that materiality ambiguity (i.e., uncertainty) results in non-uniform financial statements, allowing the auditor a great deal of latitude in determining what is material.[26] Thus, some auditors provide more leeway in determining the amount of misstatement that is material. A more lenient evaluation strategy in turn influences the manager to increase the amount of overstatement bias. However, Proposition 2 shows that when the cost of audit failure is large relative to the cost of extending audit procedures ($L > 2R$), an increase in materiality uncertainty induces a more conservative evaluation rule.

---

[25] In the knife-edge case where $L = 2R$, the auditor's evaluation strategy, c_z, and the manager's choice of overstatement bias, ω, are unaffected by materiality uncertainty, σ_m.

[26] See Jennings et al. (1985).

Based on the manager's equilibrium condition in Proposition 1, we find that materiality uncertainty σ_m affects the manager's decision only through its impact on c_z. If the auditor's evidence evaluation c_z decreases, then the probability of the auditor finding evidence of fraud increases. This induces the manager to decrease the amount of overstatement bias. Thus:

$$\frac{d\omega}{d\sigma_m} < 0. \tag{12}$$

On the other hand, a change in materiality uncertainty, σ_m, affects the auditor's choice of evaluation strategy, c_z, in two ways, directly and indirectly. The total change in the auditor's evaluation strategy is:

$$\frac{dc_z}{d\sigma_m} = \frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m} < 0. \tag{13}$$

We obtain the direct effect from the auditor's equilibrium strategy in Proposition 1. The direct effect is negative if and only if $\Phi^{-1}((L - R)/L) > 0$, where Φ is the standard normal cumulative distribution function. This condition is satisfied if and only if $L > 2R$. Hence, the direct impact of materiality uncertainty σ_m on evidence evaluation c_z is negative if and only if $L > 2R$.

The indirect effect is the impact that materiality uncertainty σ_m has on evidence evaluation c_z through its impact on overstatement bias ω. As discussed above (see Expression (12)), we know that overstatement bias ω decreases in σ_m. From Proposition 1 we know that evidence evaluation c_z decreases in intentional misstatement ω. Therefore, the indirect effect is *positive*. However the indirect effect is always overwhelmed by the direct effect because the auditor's reaction to a change in overstatement bias ω is not particularly strong relative to his reaction to σ_m. This is because the auditor is concerned with unintentional and intentional misstatements and the direct effect incorporates both types of misstatement, whereas the indirect effect incorporates only intentional misstatements. Thus the direct effect, arising from $\partial c_z / \partial \sigma_m$ drives the result of Proposition 2.

The change in sign at $L = 2R$ that we observe in Proposition 2 originates from the auditor's payoff structure.[27] The auditor incurs zero cost for accepting an immaterial misstatement, but as soon as the misstatement reaches the materiality threshold he immediately incurs an expected cost of L. Thus in choosing c_z (which also determines the threshold of acceptable updated misstatement, $\mu_{\theta|x_z=c_z,z}$) the auditor weighs the relative costs of accepting and extending procedures based on L and R. The auditor chooses an $x_z = c_z$ such that the conditional expected loss from a type II error (the constant cost, L, multiplied by the updated probability of a material misstatement) is exactly the same as the certain cost of extending audit procedures, R. For any x_z smaller than this critical value, the auditor would strictly be better off by *not* extending procedures and for any larger x_z, the auditor would be strictly better off by extending procedures. Consequently, at $x_z = c_z$ the updated probability of material misstatement is equal to the constant R/L and for every observed z, the maximum updated probability of audit failure is R/L.

---

[27] Alternative auditor payoffs lead to similar results. For example, if the auditor's expected cost of audit failure is $(\theta_z - m_z)L$ or $\theta_z L$, then the relative weightings of L and R also affect the choice of c_z where c_z decreases in L. The cost $(\theta_z - m_z)L$ can be described as Max$[0, \theta_z - m_z]$, which is continuous in θ_z. In this case, the auditor decreases c_z and the manager decreases ω for increases in materiality uncertainty regardless of the relation between L and R. However, when the auditor's cost of audit failure is $\theta_z L$, then, as in Proposition 2, the effect of materiality uncertainty depends on the relative sizes of L and R. (Details of these alternative analyses are available upon request.) This suggests that the discontinuity of the auditor's cost function at $(\theta_z - m_z)$ influences whether the size of L plays a role in how materiality uncertainty affects the equilibrium behavior of the auditor and manager.

Suppose $L = 2R$ or $R/L = 1/2$. In this case, the auditor is indifferent between acceptance and extending procedures if and only if $1 - \Phi\left(\frac{M_z - \mu_{\theta|x_z,z}}{\sqrt{(1-\alpha)\sigma_{y|z}^2 + \sigma_m^2}}\right) = 1/2$. This implies that the auditor chooses a critical value c_z that equates the threshold of updated misstatement to expected materiality, $\mu_{\theta|x_z=c_z,z} = M_z$, because $1 - \Phi(0) = 1/2$.

When the cost of audit failure is high relative to the cost of rejection ($L > 2R$), the auditor evaluates the evidence more conservatively because he has a greater incentive to avoid audit failure. Greater uncertainty about the materiality threshold increases the updated probability of material misstatement, exposing the auditor to a greater risk of audit failure. Consequently, the auditor evaluates the evidence more conservatively for higher levels of materiality uncertainty. If $L > 2R$, then $R/L < 1/2$ and the auditor is indifferent between accepting and extending procedures if and only if $\mu_{\theta|x_z=c_z,z} < M_z$. The updated probability of material misstatement is increasing in c_z and σ_m. In order to achieve a constant updated probability of material misstatement, R/L, for a given amount of overstatement bias ω at $x_z = c_z$, an increase in the exogenous parameter, σ_m, must be offset by a commensurate decrease in the auditor's strategic choice, c_z. This is the direct effect of materiality uncertainty σ_m on the auditor's evaluation strategy c_z.

On the other hand, when $L < 2R$, the auditor chooses an evaluation strategy c_z such that at $x_z = c_z$ the updated expected misstatement is *greater than* expected materiality. As a result, the updated probability of material misstatement is decreasing in materiality uncertainty for a given amount of overstatement bias. The effects of materiality uncertainty on the auditor's strategy in turn influences the manager to decrease overstatement bias when $L > 2R$ or increase overstatement bias when $L < 2R$.

Proposition 2 shows that the effect of materiality uncertainty depends on the relation between L and R. However, we anticipate that $L > 2R$ because the expected economic consequences of audit failure are usually thought to be large relative to the expected cost of extending procedures. The auditor faces a large expected cost from audit failure due to the potential losses from investors and reputation losses. As a result, increased materiality uncertainty produces a more conservative auditor evaluation strategy and a lesser amount of overstatement bias.

Materiality Effects on Audit Risk

According to SAS No. 47 (AICPA 1984, para. .02), audit risk is defined as "the risk that the auditor may unknowingly fail to appropriately modify his opinion on financial statements that are materially misstated." Audit risk in our setting is simply the average probability that the auditor accepts a materially misstated earnings report.[28] Proposition 3 describes the effects of materiality uncertainty on audit risk.

---

[28] There are two potential measures of audit risk, *ex ante* and *ex post* audit risk. Most often studies (see for example, Newman and Noel 1989; Patterson 1993) use *ex ante* audit risk (risk prior to evidence observation) because it coincides with SAS No. 47's audit risk formula that is "inherent risk times control risk times detection risk." *Ex post* audit risk depends on evidence observation and is equal to the probability in the braces of Expression (15). In our model this never exceeds R/L when the auditor accepts. To be consistent with prior studies and SAS No. 47, we use the *ex ante* measure of audit risk.

Proposition 3: When $L > 2R$, audit risk, defined as:

$$\int_{-\infty}^{c_z} (1 - H(0|x_z)) f(x_z)\, dx_z, \tag{14}$$

has three additive components resulting from an increase in materiality uncertainty.

1) The first component is the direct effect of an increase in materiality uncertainty, holding c_z and ω constant. This component increases in materiality uncertainty.
2) The second component is the indirect effect of an increase in materiality uncertainty due to its effect on c_z. This component decreases in materiality uncertainty.
3) The third component is the indirect effect of an increase in materiality uncertainty due to its effect on ω. This component either increases or decreases in materiality uncertainty.

In standard normal form, audit risk is:

$$\int_{-\infty}^{c_z} \left\{1 - \Phi\left(\frac{M_z - \mu_{\theta|x_z,z}}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)\right\} f(x_z)\, dx_z. \tag{15}$$

When $L > 2R$, the auditor's evaluation strategy and the manager's overstatement bias decrease in materiality uncertainty. The decrease in the auditor's evaluation strategy corresponds to a decrease in audit risk, while a decrease in the manager's overstatement bias may increase or decrease audit risk. On the other hand, for given values of c_z and ω, the direct effect of increasing materiality uncertainty, σ_m, is to increase audit risk. Because of these three distinct effects of materiality uncertainty on audit risk, we cannot predict its net affect on the overall change in audit risk.[29]

Figure 1 illustrates this result with a numerical example for two different earnings reports.[30] In this example, audit risk eventually decreases in σ_m, because for larger values of σ_m the rate of decrease in the auditor's evaluation strategy is higher. As a result, audit risk is decreasing at a faster rate from the indirect effects of materiality uncertainty than from the direct effects. Figure 2, Panel A shows how c_z changes as materiality uncertainty increases. Both the size of the earnings report and materiality uncertainty affect the changes in audit risk. When materiality uncertainty is high, the auditor is relatively more cautious in his evidence evaluation for subsequent increases in materiality uncertainty. This effect is more pronounced for higher earnings reports because they signal the possibility of a higher level of system error.

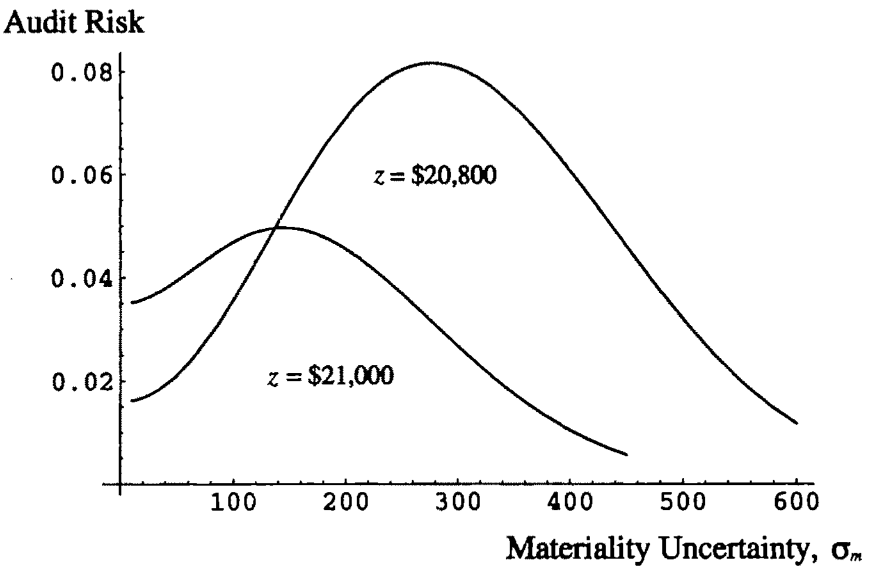
Proposition 3 also provides intuition about the effect that materiality uncertainty has on the auditor's expected payoff (see Expression (7)). The auditor's expected payoff is

---

[29] When $L < 2R$, the analytical results for changes in audit risk are also ambiguous. In this case, we cannot determine how audit risk changes due to the direct effect, but are unable to find a numerical example in which audit risk decreases in materiality uncertainty when $L < 2R$.

[30] We calculated all of our numerical examples using the following parameters: $\sigma_x = 8.41$, $\sigma_y = 164$, $\sigma_\epsilon = 280$, $\mu_y = 20{,}000$, $\pi = .05$, $p = 1.3$, $L = 1{,}000$, $R = 200$.

FIGURE 1
The Relation between Audit Risk and Materiality Uncertainty for Two Earnings Reports
$L > 2R$



Mean materiality, M_Z is .05 z.
Unconditional expected earnings, μ_y, is $20,000

equal to negative L times audit risk minus R times the probability of extending procedures. The indirect effects of materiality uncertainty affect both audit risk and the probability of extending procedures, often in opposite ways. For example the marginal decrease in audit risk times L exactly offsets the marginal increase in the probability of extending procedures times R when the auditor's evaluation rule decreases. However, the direct affect of materiality uncertainty only changes audit risk and we know that the direct effect of materiality uncertainty increases audit risk (reduces the auditor's expected payoff). While we cannot determine how the auditor's expected payoff changes due to a change in materiality uncertainty, numerical examples suggest that the auditor's expected payoff decreases in materiality uncertainty.

The Effect of the Earnings Report on the Equilibrium Strategies

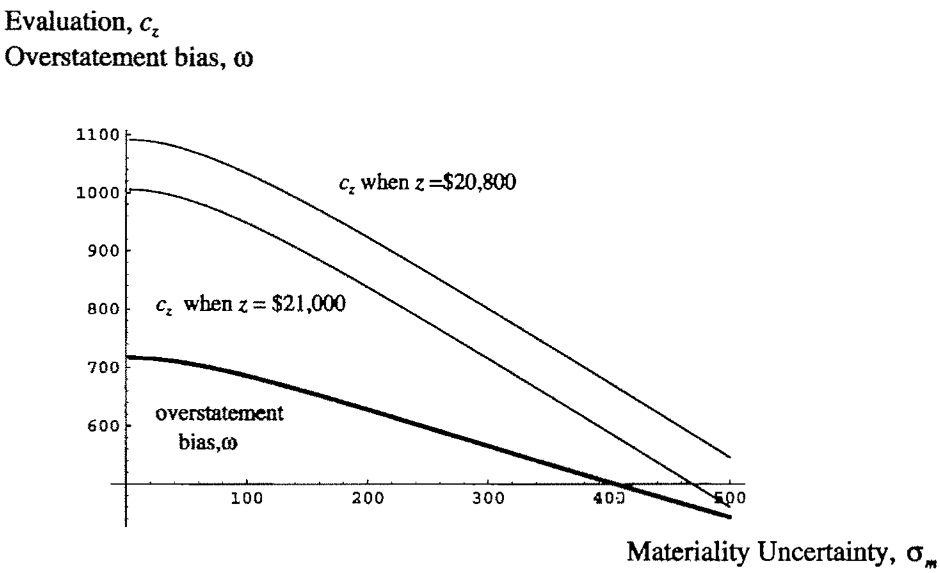
The earnings report provides the auditor with information about unintentional system error and the expected dollar amount of materiality, $M_z = \pi z$. As the amount of reported earnings increases, expected materiality increases. The professional literature suggests that as the dollar amount of materiality increases, the auditor increases his evaluation strategy.[31] However, Proposition 4 shows that in our strategic setting just the opposite may occur.

---

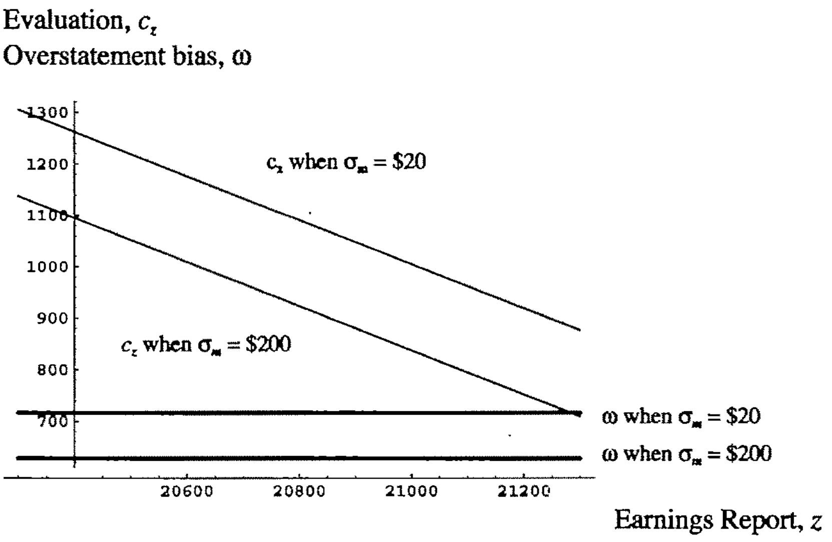
[31] See Konrath (2001).

FIGURE 2
The Relation among the Evaluation c_z, Overstatement Bias ω, Materiality Uncertainty σ_m, and the Earnings Report z
$L > 2R$

Panel A: Effects from a Change in Materiality Uncertainty



Panel B: Effects from a Change in the Earnings Report



Mean materiality, M_z is .05 z.
Unconditional expected earnings, μ_y, is \$20,000.

Proposition 4: When the earnings report, z, increases:

1) the auditor's evaluation strategy, c_z, decreases if and only if

$$\sigma_e^2 > \frac{\pi\sigma_x^2\sigma_y^2}{(1-\pi)\sigma_x^2 - \pi\sigma_y^2}$$

and $(1-\pi)\sigma_x^2 > \pi\sigma_y^2$, and

2) the manager's overstatement bias, ω, does not change.

The condition in Proposition 4 part (1) depends on the system variance σ_e^2 relative to the materiality percentage π and the variances of both audit evidence and true earnings. If the system variance is large and the materiality percentage π is small so that $(1-\pi)\sigma_x^2 > \pi\sigma_y^2$, then an increase in the earnings report increases the probability of material error. Consequently, the auditor's evaluation strategy, c_z, decreases.

Figure 2, Panel B illustrates how the evaluation strategy, c_z, changes for an increase in the earnings report, z, when the condition of Proposition 4 part (1) holds. Figure 2, Panel B also illustrates the second part of Proposition 4. The amount of the manager's overstatement bias is unaffected by the earnings report, z, because he does not observe earnings prior to his strategy choice.

V. CONCLUSION

Recently, the accounting profession and the SEC have focused attention on the issue of materiality that allows managers to overstate earnings by an immaterial amount. The SEC is concerned that the amount managers manipulate their reports depends on the materiality threshold. If misstatements fall below the materiality threshold, then the auditor is willing to accept those misstatements because, theoretically, they have no impact on the decisions of reasonable users. Furthermore, some controversy exists as to whether materiality should be quantitatively standardized. However, in considering this issue, careful attention should be given to how strategic players would react to the uncertainty of a materiality threshold.

We investigate, in a game-theoretic model, how uncertainty about the materiality threshold affects the conservatism of the auditor's evidence evaluation choice, the extent of the manager's overstatement bias, and the risk of materially misstated financial statements. We find that the auditor's conservatism increases in the uncertainty of materiality when the expected cost of audit failure is large relative to the expected cost of extending audit procedures. Auditor conservatism, in turn, induces the manager to decrease the extent of overstatement bias. Just the opposite occurs when the expected cost of extending procedures is large relative to the expected cost of audit failure. However, because the expected cost of audit failure is likely to be large relative to the expected cost of extending audit procedures, our model suggests that a fixed materiality standard would be detrimental in promoting auditor avoidance of audit failure.

Surprisingly, though, audit risk or the average probability that audit failure occurs, can actually increase as the uncertain materiality threshold increases even though the auditor evaluates the evidence more conservatively and the extent of intentional overstatement decreases. The change in the auditor's equilibrium strategy decreases audit risk indirectly. However, the variance of the materiality threshold directly increases audit risk and this increase can have a dominant effect on the overall change in audit risk.

Of course, our study is limited by our stylized setting. Typically decisions about materiality are more complex, auditors' payoffs are affected by additional legal concerns, and managers consider a vast array of motivational factors. However, adding more complexity to the model is unlikely to negate the result that materiality uncertainty can decrease the extent of overstatement bias while simultaneously increasing or decreasing audit risk.

APPENDIX

Proof of Proposition 1

Part (1)

In equilibrium, $c_z = x_z$ equates the left-hand and right-hand sides of Expression (9) so that:

$$\mu_{\theta|x_z,z} = M_z - \Phi^{-1}\left(\frac{L-R}{L}\right)\sqrt{(1-\alpha)\sigma_{y|z}^2 + \sigma_m^2}.$$

Or:

$$\begin{aligned}&\alpha x_z + ((1-\alpha)(1-\delta) - \pi)z + (1-\alpha)(\delta\omega - (1-\delta)\mu_y)\\ &\quad = -\Phi^{-1}\left(\frac{L-R}{L}\right)\sqrt{(1-\alpha)\sigma_{y|z}^2 + \sigma_m^2}\end{aligned}$$

$$\begin{aligned}\Leftrightarrow \alpha x_z + ((1-\alpha)(1-\delta) - \pi)z = &-(1-\alpha)(\delta\omega - (1-\delta)\mu_y)\\ &- \Phi^{-1}\left(\frac{L-R}{L}\right)\sqrt{(1-\alpha)\sigma_{y|z}^2 + \sigma_m^2}\end{aligned}$$

$$\begin{aligned}\Leftrightarrow x_z &= \frac{(\pi - (1-\alpha)(1-\delta))z - (1-\alpha)(\delta\omega - (1-\delta)\mu_y) - \Phi^{-1}\left(\frac{L-R}{L}\right)\sqrt{(1-\alpha)\sigma_{y|z}^2 + \sigma_m^2}}{\alpha}\\ &= \frac{(\pi - (1-\alpha)(1-\delta))}{\alpha}z + \frac{-(1-\alpha)(\delta\omega - (1-\delta)\mu_y) - \Phi^{-1}\left(\frac{L-R}{L}\right)\sqrt{(1-\alpha)\sigma_{y|z}^2 + \sigma_m^2}}{\alpha}.\end{aligned}$$

Thus $c_z = az + b$ where:

$$a = \frac{(\pi - (1-\alpha)(1-\delta))}{\alpha} \quad \text{and}$$

$$b = \frac{-(1-\alpha)(\delta\omega - (1-\delta)\mu_y) - \Phi^{-1}\left(\frac{L-R}{L}\right)\sqrt{(1-\alpha)\sigma_{y|z}^2 + \sigma_m^2}}{\alpha}.$$

Part (2)

Recall from Expression (8) that the manager's expected payoff is:

$$U = \int_{-\infty}^{\infty} \{z - p\omega(1 - F(c_z))\}\lambda(z)\,dz.$$

$$= \mu_y + \omega - p\omega + p\omega \int_{-\infty}^{\infty} F(c_z)\lambda(z)\,dz \tag{A1}$$

where:

$$\int_{-\infty}^{\infty} F(c_z)\lambda(z)\,dz = \int_{-\infty}^{\infty} \text{Prob}(x_z < c_z|z)\lambda(z)\,dz.$$

We also know from general probability theory (see Ash [1972, 240–243] for discussion) that Prob($\{z \in A, x \in B\}$) = $\int_{-\infty}^{\infty}$ Prob($x_z < c_z|z)\lambda(z)\,dz$ where event B is $x_z < c_z$ and event A is $z \in \{-\infty,\infty\}$. Furthermore, because all auditor equilibrium evaluation rules c_z have the general form $c_z = az + b$, $x_z < c_z$ is equivalent to $x_z < az + b$ and Prob($\{z \in A, x \in B\}$) = Prob($x - az < b$).

Therefore, because the random variable $x - az$ is normal:

$$\text{Prob}(x - az < b) = \Phi\left(\frac{b - (E[x - az])}{\sqrt{\text{Var}[x - az]}}\right)$$

where:

$$E[x - az] = E[e + \omega + \xi - a(y + e + \omega)] = \omega - a(\mu_y + \omega) \quad \text{and}$$

$$\text{Var}[x - az] = \text{Var}[(1 - a)e - ay + \xi] = (1 - a)^2\sigma_e + a^2\sigma_y^{\ 2} + \sigma_x^{\ 2}.$$

Thus:

$$\int_{-\infty}^{\infty} F(c_z)\lambda(z)\,dz = \Phi\left(\frac{b - (\omega - a(\mu_y + \omega))}{\sqrt{(1 - a)^2\sigma_e^{\ 2} + a^2\sigma_y^{\ 2} + \sigma_x^{\ 2}}}\right)$$

$$= \Phi\left(\frac{b + a(\mu_y + \omega) - \omega}{\sqrt{(1 - a)^2\sigma_e^{\ 2} + a^2\sigma_y^{\ 2} + \sigma_x^{\ 2}}}\right)$$

and the manager's expected payoff in Expression (A1) can be written as:

$$U = \mu_y + \omega - p\omega(1 - \mathcal{F}(c)) \tag{A2}$$

where $c = a(\mu_y + \omega) + b$ and $\mathcal{F}(\cdot)$ is the unconditional distribution function of x with mean ω and variance $(1 - a)^2\sigma_e^{\ 2} + a^2\sigma_y^{\ 2} + \sigma_x^{\ 2}$.

Consequently, using the chain rule to differentiate U in Expression (A2), we obtain:

$$\frac{dU}{d\omega} = 1 - p(1 - \mathcal{F}(c)) - (1 - a)p\omega f(c) = 0 \tag{A3}$$

as the manager's equilibrium condition where a and b are defined as in part (1).

Proof that ω Satisfying Expression (A3) Is a Payoff Maximum

$$\text{Let } U_\omega = \frac{dU}{d\omega} \quad \text{and} \quad U_{\omega j} = \frac{d^2U}{d\omega\, dj}.$$

Recall that $c = a(\omega + \mu_y) + b$. Thus, $c - \omega = (a - 1)\omega + a\mu_y + b$.

Denote $\phi(\cdot)$ and $\Phi(\cdot)$ as the standard normal density and cumulative distribution, respectively. Using the chain rule where:

$$\mathfrak{f}(c) = \frac{1}{(1 - a)^2\sigma_e^2 + a^2\sigma_y^2 + \sigma_x^2}\,\phi\left(\frac{c - \omega}{(1 - a)^2\sigma_e^2 + a^2\sigma_y^2 + \sigma_x^2}\right)$$

and

$$\mathfrak{F}(c) = \Phi\left(\frac{c - \omega}{(1 - a)^2\sigma_e^2 + a^2\sigma_y^2 + \sigma_x^2}\right),$$

we derive the manager's second order condition:

$$\begin{aligned} U_{\omega\omega} &= -(1 - a)p\mathfrak{f}(c) - (1 - a)p\mathfrak{f}(c) \\ &\quad - (1 - a)^2\, p\omega\mathfrak{f}(c)\,\frac{c - \omega}{(1 - a)^2\sigma_e^2 + a^2\sigma_y^2 + \sigma_x^2} \\ &= -(1 - a)\left(2p\mathfrak{f}(c) + (1 - a)p\omega\mathfrak{f}(c)\,\frac{c - \omega}{(1 - a)^2\sigma_e^2 + a^2\sigma_y^2 + \sigma_x^2}\right) \\ &= -(1 - a)p\mathfrak{f}(c)\left(2 + (1 - a)\omega\,\frac{c - \omega}{(1 - a)^2\sigma_e^2 + a^2\sigma_y^2 + \sigma_x^2}\right). \end{aligned} \tag{A4}$$

Next, fix ω at $\hat{\omega}$ that solves $U_\omega = 0$ and consider the limiting behavior of $U_\omega(b)$ as b decreases and increases. Our objective is to determine the sign of $U_{\omega\omega}$ when $U_\omega = 0$ and $\omega = \hat{\omega}$. We determine this sign by first determining the sign of $U_{\omega b}$ and then exploit the fact (that we will show later) that $U_{\omega\omega} = -(1 - a)\{U_{\omega b} + p\mathfrak{f}(c)\}$. In other words, if $U_{\omega b}$ is positive, then $U_{\omega\omega}$ is negative if and only if $1 > a$.

$$\begin{aligned} \underset{b\to-\infty}{\text{Limit}}[U_\omega] &= \underset{b\to-\infty}{\text{Limit}}[1 - p(1 - \mathfrak{F}(c)) - (1 - a)p\omega\mathfrak{f}(c)] \\ &= 1 - p < 0 \text{ because } 1 < p. \end{aligned}$$

Also

$$\begin{aligned} \underset{b\to\infty}{\text{Limit}}[U_\omega] &= \underset{b\to\infty}{\text{Limit}}[1 - p(1 - \mathfrak{F}(c)) - (1 - a)p\omega\mathfrak{f}(c)] \\ &= 1 > 0. \end{aligned}$$

Consequently, as b increases, U_ω is increasing in b the first time it crosses zero. Furthermore:

$$U_{\omega b} = pf(c) + (1-a)p\omega f(c)\frac{c-\omega}{(1-a)^2\sigma_e^2 + a^2\sigma_y^2 + \sigma_x^2}$$

$$= pf(c)\left\{1 + (1-a)\omega\frac{c-\omega}{(1-a)^2\sigma_e^2 + a^2\sigma_y^2 + \sigma_x^2}\right\}. \tag{A5}$$

Now the term in braces in (A5) remains positive if $(1 - a) > 0$ because it increases in b. If $(1 - a) < 0$, then the term in brackets in (A5) can be zero for only one value of b and we know that at $b = +\infty$, $U_{\omega b} > 0$. Thus $U_{\omega b}$ can cross zero only once and must be positive when $U_\omega = 0$. As a result, $U_{\omega b} > 0$ whenever $U_\omega = 0$.

We substitute $U_{\omega b}$ (Expression (A5)) into $U_{\omega\omega}$ (Expression (A4)) to obtain:

$$U_{\omega\omega} = -(1-a)\{U_{\omega b} + pf(c)\} \quad \text{when } U_\omega = 0. \tag{A6}$$

Thus:

$$U_{\omega\omega} < 0 \Leftrightarrow (1-a) > 0 \quad \text{when } U_\omega = 0.$$

Finally:

$$1 - a = 1 - \frac{\pi - (1-\alpha)(1-\delta)}{\alpha} = 1 - \frac{\pi - \dfrac{\sigma_x^2}{\sigma_{y|z}^2 + \sigma_x^2}\dfrac{\sigma_e^2}{\sigma_e^2 + \sigma_y^2}}{\dfrac{\sigma_{y|z}^2}{\sigma_{y|z}^2 + \sigma_x^2}}$$

$$= 1 - \frac{\pi(\sigma_{y|z}^2 + \sigma_x^2) - \dfrac{\sigma_x^2\sigma_e^2}{\sigma_e^2 + \sigma_y^2}}{\sigma_{y|z}^2} = 1 - \frac{\pi\left(\dfrac{\sigma_y^2\sigma_e^2}{\sigma_e^2 + \sigma_y^2} + \sigma_x^2\right) - \dfrac{\sigma_x^2\sigma_e^2}{\sigma_e^2 + \sigma_y^2}}{\dfrac{\sigma_y^2\sigma_e^2}{\sigma_e^2 + \sigma_y^2}}$$

$$= 1 - \frac{\pi(\sigma_y^2\sigma_e^2 + \sigma_x^2\sigma_e^2 + \sigma_x^2\,\sigma_y^2) - \sigma_x^2\sigma_e^2}{\sigma_y^2\sigma_e^2} = (1-\pi)\left(1 - \frac{\sigma_x^2}{\sigma_y^2}\right) - \pi\frac{\sigma_x^2}{\sigma_e^2}$$

$$\Rightarrow (1-a) > 0 \Leftrightarrow \sigma_e^2 > \frac{\pi\sigma_x^2\sigma_y^2}{(1-\pi)(\sigma_x^2 + \sigma_y^2)}.$$

which is true based on our assumption of a noisy information system.

Thus, ω is a unique maximizing strategy for the manager.

Proof of Equilibrium Uniqueness

In order to prove the equilibrium $\{\omega, c_z\}$ is unique, we must show three things.

(1) There exists a unique strategy that maximizes the auditor's expected payoff given any choice of ω by the manager.
This is satisfied based on part (1) of Proposition 1, which shows that c_z is linear in ω.

(2) There exists a unique strategy that maximizes the manager's expected payoff given any choice of $c_z = az + b$ by the auditor.
This is satisfied by the proof above that shows $U_{\omega\omega} < 0$, whenever $U_\omega = 0$.

(3) There is only one set of strategies $(\{\omega, c_z\})$ that satisfy (1) and (2).
To show this we consider the manager's equilibrium condition.
Let $EqU = 1 - p(1 - \mathfrak{F}(c)) - (1 - a)p\omega f(c) = 0$ where a and b are defined as in part (1) of Proposition 1. Thus, $\frac{dEqU}{d\omega} = U_{\omega\omega} + U_{\omega b}\frac{db}{d\omega}$. Because $\frac{db}{d\omega} = \frac{-(1-\alpha)\,\delta}{\alpha} < 0$, $U_{\omega\omega} < 0$ and $U_{\omega b} > 0$ for the equilibrium choice of ω, $\frac{dEqU}{d\omega} < 0$. Thus the equilibrium of our game is unique.

Proof of Proposition 2

By implicit differentiation:

$$\frac{d\omega}{d\sigma_m} = -\frac{\frac{dEqU}{d\sigma_m}}{\frac{dEqU}{d\omega}}$$

and in Proposition 1 we show that $\frac{dEqU}{d\omega} = < 0$. Using the chain rule we obtain:

$$\frac{dc_z}{d\sigma_m} = \frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m}.$$

Part (1)

$$\frac{dEqU}{d\sigma_m} = U_{\omega b}\frac{db}{d\sigma_m} = -U_{\omega b}\left(\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right) < 0$$

$$\Leftrightarrow \frac{L-R}{L} > 1/2 \quad \text{or} \quad 2R < L.$$

Consequently, $\frac{d\omega}{d\sigma_m} < 0 \Leftrightarrow 2R < L.$

Next $\frac{\partial c_z}{\partial \sigma_m} = -\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}$, $\frac{\partial c_z}{\partial \omega} = -\frac{(1-\alpha)\delta}{\alpha}$ and

$$\frac{d\omega}{d\sigma_m} = -\frac{-U_{\omega b}\left(\dfrac{\Phi^{-1}\left(\dfrac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}^2+\sigma_m^2}}\right)}{U_{\omega\omega}+U_{\omega b}\dfrac{db}{d\omega}}$$

$$= \frac{-U_{\omega b}\left(\dfrac{\Phi^{-1}\left(\dfrac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}^2+\sigma_m^2}}\right)}{(1-a)\{U_{\omega b}+pf(c)\}+U_{\omega b}\dfrac{(1-\alpha)\delta}{\alpha}}.$$

Thus:

$$\frac{\partial c_z}{\partial\omega}\frac{d\omega}{d\sigma_m} = -\frac{(1-\alpha)\delta}{\alpha}\,\frac{-U_{\omega b}\left(\dfrac{\Phi^{-1}\left(\dfrac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}^2+\sigma_m^2}}\right)}{(1-a)\{U_{\omega b}+pf(c)\}+U_{\omega b}\dfrac{(1-\alpha)\delta}{\alpha}}$$

$$< \frac{(1-\alpha)\delta}{\alpha}\,\frac{\left(\dfrac{\Phi^{-1}\left(\dfrac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}^2+\sigma_m^2}}\right)}{\dfrac{(1-\delta)-\pi+\alpha\delta+(1-\alpha)\delta}{\alpha}}$$

when $L > 2R$. This implies that:

$$\frac{dc_z}{d\sigma_m} = \frac{\partial c_z}{\partial\sigma_m}+\frac{\partial c_z}{\partial\omega}\frac{d\omega}{d\sigma_m} < -\frac{\Phi^{-1}\left(\dfrac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}^2+\sigma_m^2}} + \frac{(1-\alpha)\delta}{\alpha}\left(\frac{\Phi^{-1}\left(\dfrac{L-R}{L}\right)\sigma_m}{(1-\pi)\sqrt{(1-\alpha)\sigma_{y|z}^2+\sigma_m^2}}\right)$$

$$= -\frac{\Phi^{-1}\left(\dfrac{L-R}{L}\right)\sigma_m}{\sqrt{(1-\alpha)\sigma_{yz}^2+\sigma_m^2}}\,\frac{(1-\pi-(1-\alpha)\delta)}{\alpha(1-\pi)}$$

$$= -\frac{\Phi^{-1}\left(\dfrac{L-R}{L}\right)\sigma_m}{\sqrt{(1-\alpha)\sigma_{y|z}^2+\sigma_m^2}}\,\frac{(1-a)}{(1-\pi)} < 0 \text{ when } L > 2R.$$

To show that $\frac{dc_z}{d\sigma_m} < 0$ implies $L > 2R$, suppose by way of contradiction that $L \leq 2R$. Then by the proofs of parts (2) and (3) we arrive at a contradiction. Thus $\frac{dc_z}{d\sigma_m} < 0 \Leftrightarrow L > 2R$.

Part (2)

Also:

$$\frac{dEqU}{d\sigma_m} = U_{\omega b}\frac{db}{d\sigma_m} = -U_{\omega b}\left(\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right) > 0$$
$$\Leftrightarrow \frac{L-R}{L} < 1/2 \text{ or } 2R > L.$$

Consequently, $\frac{d\omega}{d\sigma_m} > 0 \Leftrightarrow 2R > L$.

The proof of the sign of $\frac{dc_z}{d\sigma_m}$ when $L < 2R$ is similar to part (1) except:

$$\frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m} = -\frac{(1-\alpha)\delta}{\alpha}\frac{-U_{\omega b}\left(\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)}{(1-a)\{U_{\omega b} + pf(c)\} + U_{\omega b}\frac{(1-\alpha)\delta}{\alpha}}$$
$$> \frac{(1-\alpha)\delta}{\alpha}\left(\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{(1-\pi)\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)$$

when $L < 2R$.

Thus $\frac{dc_z}{d\sigma_m} = \frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m} > -\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\frac{(1-a)}{(1-\pi)} > 0$ when $L < 2R$.

To show that $\frac{dc_z}{d\sigma_m} > 0$ implies $L < 2R$, suppose by way of contradiction that $L \geq 2R$. Then by the proofs of parts (1) and (3) we arrive at a contradiction. Thus $\frac{dc_z}{d\sigma_m} > 0 \Leftrightarrow L < 2R$.

Part (3)
The Knife-Edge Case, L = 2R.

$$\frac{dEqU}{d\sigma_m} = U_{\omega b}\frac{db}{d\sigma_m} = -U_{\omega b}\left(\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right) = 0$$

$$\Leftrightarrow \frac{L-R}{L} = 1/2 \quad \text{or} \quad 2R = L.$$

Consequently, $\frac{d\omega}{d\sigma_m} = 0 \Leftrightarrow 2R = L.$

The proof for the sign of $\frac{dc_z}{d\sigma_m} = 0$ when $L = 2R$ is similar to part (1) except:

$$\frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m} = -\frac{(1-\alpha)\delta}{\alpha}\frac{-U_{\omega b}\left(\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)}{(1-a)\{U_{\omega b} + pf(c)\} + U_{\omega b}\frac{(1-\alpha)\,\delta}{\alpha}} = 0$$

when $L = 2R$.

Thus $\frac{dc_z}{d\sigma_m} = \frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m} = \frac{\partial c_z}{\partial \sigma_m} = -\frac{\Phi^{-1}\left(\frac{L-R}{L}\right)\sigma_m}{\alpha\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}} = 0$ when $L = 2R$.

To show that $\frac{dc_z}{d\sigma_m} = 0$ implies $L = 2R$, suppose by way of contradiction that $L \neq 2R$. Then by the proofs of parts (1) and (2) we arrive at a contradiction. Thus $\frac{dc_z}{d\sigma_m} = 0 \Leftrightarrow L = 2R$.

Proof of Proposition 3

Let Audit Risk be denoted as:

$$AR = \int_{-\infty}^{c_z}\left\{1 - \Phi\left(\frac{M_z - \mu_{\theta|x_z,z}}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)\right\} f(x_z)\,dx_z.$$

The overall change in audit risk due to a change in materiality uncertainty is:

$$\frac{dAR}{d\sigma_m} = \frac{\partial AR}{\partial c_z}\left\{\frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m}\right\} + \frac{\partial AR}{\partial \omega}\frac{d\omega}{d\sigma_m} + \frac{\partial AR}{\partial \sigma_m}$$

where $\frac{\partial AR}{\partial c_z}\left\{\frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m}\right\}$ is an indirect change arising from the corresponding change in c_z, $\frac{\partial AR}{\partial \omega}\frac{d\omega}{d\sigma_m}$ is an indirect change arising from a corresponding change in ω, and $\frac{\partial AR}{\partial \sigma_m}$ is the direct change in audit risk.

Part (1)

$$\frac{\partial AR}{\partial \sigma_m}$$

$$= \int_{-\infty}^{c_z} f(x_z)\left\{\phi\left(\frac{M_z - \mu_{\theta|x_z,z}}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)\frac{(M_z - \mu_{\theta|x_z,z})}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\frac{\sigma_m}{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}\right\} dx_z > 0$$

because when $L > 2R$, $\frac{(M_z - \mu_{\theta|x_z,z})}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}} > 0$ for all $x_z \leq c_z$.

Part (2)

$$\frac{\partial AR}{\partial c_z}\left\{\frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m}\right\}$$

$$= f(c_z)\left\{1 - \Phi\left(\frac{M_z - \mu_{\theta|c_z,z}}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)\right\}\left\{\frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m}\right\}$$

$$= f(c_z)\frac{R}{L}\left\{\frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m}\right\} < 0$$

because in Proposition 2 we show that $\left\{\frac{\partial c_z}{\partial \sigma_m} + \frac{\partial c_z}{\partial \omega}\frac{d\omega}{d\sigma_m}\right\} < 0$ when $L > 2R$ and because at $x_z = c_z$, Expression (6) is an equality.

Part (3)

$$\frac{\partial AR}{\partial \omega}\frac{d\omega}{d\sigma_m}$$

$$= \int_{-\infty}^{c_z} f(x_z)\left\{-\phi\left(\frac{M_z - \mu_{\theta|x_z,z}}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)\right\} dx_z \frac{(-(1-\alpha)\delta)}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\frac{d\omega}{d\sigma_m}$$

$$+ \int_{-\infty}^{c_z} f(x_z)\left\{\phi\left(\frac{M_z - \mu_{\theta|x_z,z}}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)\right\} dx_z \frac{\alpha\delta}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\frac{d\omega}{d\sigma_m}$$

$$- f(c_z)\frac{R}{L}\frac{d\omega}{d\sigma_m}$$

$$= \int_{-\infty}^{c_z} f(x_z)\left\{\phi\left(\frac{M_z - \mu_{\theta|x_z,z}}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\right)\right\} dx_z \frac{\delta}{\sqrt{(1-\alpha)\sigma_{y|z}{}^2 + \sigma_m{}^2}}\frac{d\omega}{d\sigma_m}$$

$$- f(c_z)\frac{R}{L}\frac{d\omega}{d\sigma_m} < \text{or} > 0.$$

Proof of Proposition 4
Part (1) [Part (2) follows directly from Proposition 1]

Recall from Proposition 1 part (1) that $c_z = az + b$. Thus $\frac{dc_z}{dz} = a$.

From Proposition 1 we know that:

$$a = \frac{\pi(\sigma_y^2\sigma_e^2 + \sigma_x^2\sigma_e^2 + \sigma_x^2\,\sigma_y^2) - \sigma_x^2\sigma_e^2}{\sigma_y^2\sigma_e^2}.$$

The numerator of a equals:

$$\pi\sigma_y^2\sigma_e^2 + \pi\sigma_x^2\sigma_y^2 - (1 - \pi)\sigma_x^2\sigma_e^2 = \sigma_e^2(\pi\sigma_y^2 - (1 - \pi)\sigma_x^2) + \pi\sigma_x^2\sigma_y^2.$$

Consequently:

$$a < 0 \Leftrightarrow \sigma_e^2 > \frac{\pi\sigma_x^2\sigma_y^2}{(1 - \pi)\sigma_x^2 - \pi\sigma_y^2}$$

and $(1 - \pi)\sigma_x^2 - \pi\sigma_y^2 > 0$.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 1984. *Materiality and Audit Risk.* SAS No. 47. New York, NY: AICPA.

———. 1997. *Considering Fraud in a Financial Statement Audit.* SAS No. 82. New York, NY: AICPA.

———. 2001. *AICPA Professional Standards.* Volume 1. New York, NY: AICPA.

Ash, R. 1972. *Real Analysis and Probability.* New York, NY: Academic Press Inc.

Berger, J. O. 1985. *Statistical Decision Theory and Bayesian Analysis.* New York, NY: Springer-Verlag.

Engineering News-Record (ENR). 2002. Anti-fraud bill signed into law. Volume 249 (6).

Fedders, J. 1998. Qualitative materiality: The birth, struggles, and demise of an unworkable standard. *The Catholic University Law Review* (Fall): 41–91.

Fudenberg, D., and J. Tirole. 1995. *Game Theory.* Cambridge, MA: MIT Press.

Jennings, M., P. Reckers, and D. Kneer. 1985. A source of insecurity: A discussion and an empirical examination of standards of disclosure and levels of materiality in financial statements. *The Journal of Corporation Law* 10 (3): 639–688.

Johnstone, D. 1985. Statistically incoherent hypothesis tests in auditing. *Auditing: A Journal of Practice & Theory* (Fall): 156–175.

Journal of Accountancy. 2000. New SASs address communications and adjustments. Volume 189 (March): 17–18. Available at: http://www.aicpa.org/pubs/jofa/mar2000/newsl.htm.

Kiernan, J. S., and L. L. Lewin. 1994. Those monstrous damage awards against accountants. *The CPA Journal* (April): 36–41.

Kinney, W. 1975. A decision theory approach to the sampling problem in auditing. *Journal of Accounting Research* 13 (1): 117–132.

Konrath, L. F. 2001. *Auditing, A Risk Analysis Approach.* Cincinnati, OH: South-Western college Publishing.

Newman, P., and J. Noel. 1989. Error rates, detection rates, and payoff functions in auditing. *Auditing: A Journal of Practice & Theory* (Supplement): 50–63.

Newman, P., S. Rhoades, and R. Smith. 1996. Allocating audit resources to detect fraud. *Review of Accounting Studies* 1 (2): 161–182.

Patterson, E. R. 1993. Strategic sample size choice in auditing. *Journal of Accounting Research* 31 (2): 272–293.

Public Oversight Board (POB). 2000. *The Panel on Audit Effectiveness: Report and Recommendations.* Stamford, CT: POB.

Ricchiute, D. 2002. *Auditing & Assurance Services.* Cincinnati, OH: South-Western College Publishing.

Securities and Exchange Commission (SEC). 1999. *Materiality.* SEC Staff Accounting Bulletin No. 99. Washington, D.C.: Government Printing Office.

Trager, M. H. 2002. The Enron affair: From an accountant's point of view. *The Secured Lender* (58): 72–74.

Wall Street Journal. 1998. SEC readies new companies about what is "material" for disclosure. Dow Jones, Inc. (November 3): A2.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 847–874

Corporate Tax-Planning Effectiveness: The Role of Compensation-Based Incentives

John D. Phillips
University of Connecticut

ABSTRACT: This study investigates whether compensating chief executive officers and business-unit managers using after-tax accounting-based performance measures leads to lower effective tax rates, the empirical surrogate used for tax-planning effectiveness. Utilizing proprietary compensation data obtained in a survey of corporate executives, the relation between effective tax rates and after-tax performance measures is modeled and estimated using a two-step approach that corrects for the endogeneity bias associated with firms' decisions to compensate managers on a pre- versus after-tax basis. The results are consistent with the hypothesis that compensating business-unit managers, but not chief executive officers, on an after-tax basis leads to lower effective tax rates.

Keywords: *tax planning; performance measures; endogenous treatment effects.*

Data Availability: *Compustat data are publicly available. Survey data are proprietary in nature.*

I. INTRODUCTION

Effective tax planning, defined by Scholes et al. (2002) as tax planning that maximizes the firm's expected discounted after-tax cash flows, requires managers to consider their decisions' after-tax consequences. In this paper, I investigate whether after-tax accounting-based performance measures lead to lower effective tax rates (ETRs), my

---

I thank the following members of my dissertation committee for their help and guidance: Dan Collins (chairman), Murray Barrick, Ed Maydew, Forrest Nelson, Mort Pincus, and Rick Tubbs. I also appreciate the helpful comments and suggestions of Ramji Balakrishnan, Amy Dunbar, Bill Greene, Bruce Johnson, Lil Mills, Sonja Olhoft Rego, Tom Omer, George Plesko, Richard Sansing, Peter Wilson, and workshop participants at the University of Connecticut, The University of Iowa, The University of Texas at Austin, 1999 University of Illinois Tax Symposium, 2000 American Taxation Association Midyear Meeting, and the 2000 American Accounting Association Northeast Regional Meeting, along with two anonymous reviewers and the associate editor. I am indebted to the Financial Executives Institute, particularly Carol Garnant and John Porter, for sponsoring the survey. I also thank Maggie Jesse and Josh Kaine for their help in implementing the survey and Kristi Herbrechtsmeyer for her help in collecting data. The financial assistance of The University of Iowa Tippie College of Business, the University of Connecticut School of Business, and the Deloitte & Touche Foundation is gratefully acknowledged. This paper is based on my dissertation completed at The University of Iowa.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted April 2001
Accepted February 2003

empirical surrogate for tax planning effectiveness.[1] The ETR, an income-statement-based outcome measure calculated as the ratio of total income tax expense to pre-tax income, generally measures the effectiveness of tax reduction strategies that lead to higher after-tax income. A lower ETR, however, can only proxy for tax savings and does not always imply that after-tax income and/or cash flows have been maximized.[2] Despite this limitation, the ETR has been used to measure the effectiveness of spending on the tax function (Mills et al. 1998) and corporate tax department performance (Douglas et al. 1996). Also, lowering the ETR is frequently cited as a way to increase earnings (e.g., Ziegler 1997) and increase share price (e.g., Mintz 1999; Swenson 1999).[3]

Accounting research has addressed the relation between accounting-based compensation and managers' actions (e.g., Larcker 1983; Healy 1985; Wallace 1997). This paper is the first to address whether after-tax accounting-based performance measures motivate managers to take actions that help lower their firms' ETRs and does so at both the chief executive officer (CEO) and business-unit (BU) manager levels. Prior after-tax performance measure research has focused only on the determinants of compensating CEOs using pre- versus after-tax earnings (e.g., Newman 1989; Carnes and Guffey 2000; Atwood et al. 1998; Dhaliwal et al. 2000) and provides no evidence concerning after-tax compensation's effectiveness in lowering a firm's tax liability. Extending this investigation to the BU level is motivated out of the apparent conflict between arguments that taxes should be allocated to BUs for incentive compensation purposes (e.g., McLemore 1997) with empirical observations that a majority of firms do not do so (e.g., Douglas et al. 1996).[4] The current investigation provides evidence concerning the incremental effectiveness of explicitly motivating CEOs and BU managers to incorporate tax consequences into their operating and investment decisions.

A common issue in cross-sectional studies that attempt to link a particular management accounting choice to an outcome measure is that all sample firms may be optimizing with respect to the choice being investigated (Ittner and Larcker 2001). Without addressing the endogeneity of a firm's choice, it is difficult to provide evidence consistent with this choice leading to an improved outcome. To address this issue, the relation between ETRs and CEO and BU-manager after-tax performance measures is estimated using a two-step approach that helps correct for the potential endogeneity bias associated with these two choice variables. As a first step in implementing this approach, the Antle and Demski (1988) controllability principle is used to model a firm's decisions to adopt after-tax CEO and BU-manager performance measures. To include a particular measure in a manager's compensation contract, this principle requires that the expected benefits from holding a manager responsible for a measure must be greater than the additional wage that must be paid to compensate the manager for the resulting additional risk and effort. Accordingly, an after-tax performance measure should be used as a contracting variable in a manager's incentive compensation contract only if the manager's involvement in tax-planning efforts leads to a difference between pre-tax and after-tax accounting results, which is generally reflected in the ETR. Consistent with prior research, the pre- versus after-tax CEO and BU-manager selection models include variables that control for a firm's tax-planning opportunities because the presence of such opportunities reflect the extent to which a manager's actions can be expected to lower the ETR.

---

[1] The difference in after-tax income between firms compensating its managers on a pre- versus after-tax basis is the preferable empirical surrogate for measuring the effectiveness of after-tax accounting-based compensation schemes. Using this measure is infeasible, however, due to the many unidentifiable and uncontrollable factors that affect a firm's profitability.

[2] ETR minimization does not imply a firm has engaged in effective tax planning. An extreme example where ETR minimization does not imply effective tax planning is when a firm invests solely in assets that generate a tax-free return (e.g., municipal bonds) and achieves a zero ETR when greater after-tax income and cash flows could have been achieved through investments in taxable assets. Managers attempting to maximize after-tax accounting profits, however, would not engage in this activity.

[3] Lev and Thiagarajan (1993) and Abarbanell and Bushee (1997) both provide evidence that ETRs are negatively related to future changes in earnings and future stock returns. Guenther and Jones (2002) provide evidence that ETR changes are positively related to stock returns, but only if the ETR change increases earnings. These studies suggest that a firm's ETR-lowering strategies may not always lead to a higher stock price.

[4] Thirty-two percent of this study's sample compensates its BU managers on an after-tax basis.

Even if a manager's efforts are expected to lead to a lower ETR, a firm will use an after-tax performance measure only if the expected benefits exceed the expected costs of doing so. An after-tax performance measure is expected to lead to a lower ETR because it motivates the manager's increased cooperation with tax professionals to help identify, develop, and execute tax-planning strategies. McLemore (1997, 1) cites Hewlett Packard's tax director to support the need for BU-manager involvement in tax-planning efforts:

> Tax planning is only as good as being involved in the early stages of such things as business planning, strategic planning, and merger and acquisition work....Your tax department has to be represented at the table when those decisions are made. The evolving model for the future is the tight integration of tax people with business unit planning.

Costs associated with using after-tax performance measures include the additional wage that must be paid to compensate the manager for the increased risk due to potential tax law changes and the increased effort that results from including income tax expense in the compensation contract. Other potential costs associated with after-tax compensation include the administrative cost of allocating tax expense to a firm's BUs, increased tax examination costs, and increased tax authority scrutiny. Contrary to measuring after-tax compensation's benefits via observed ETRs, there are no clear empirical surrogates for after-tax performance measures' costs. This study thus focuses on the realized benefits of compensating managers on an after-tax basis but does not provide evidence of the associated costs' magnitude.

Proprietary data obtained in a survey of corporate executives are used to construct certain test variables, including those indicating whether CEOs and BU managers are compensated using after-tax accounting-based performance measures. Publicly available data are used to construct ETRs and other test variables. The results are consistent with the hypothesis that compensating BU managers, but not CEOs, on an after-tax basis leads to lower ETRs, resulting in an estimated median tax savings of $13.3 million annually. Sensitivity tests performed on a subsample of firms with high simulated MTRs (Graham 1996) provide further evidence that low-MTR firms' potential ETR-lowering actions that could have ambiguous effects on cash flows and after-tax profits are not driving this result. Further sensitivity tests help rule out the proportion of tax function outsourcing as an alternative explanation for the statistically and economically significant negative relation between after-tax BU-manager compensation and ETRs.

The results contribute to the accounting-based compensation literature by linking after-tax accounting-based performance measures to BU-manager involvement that is incrementally effective in lowering firms' ETRs. Consistent with Guidry et al. (1999) who document bonus-induced earnings management at the BU level, this finding provides additional insight into the effect that BU-manager accounting-based incentives have on managers' actions. Also, the estimated explicit tax savings resulting from after-tax performance measures provide corporate decision makers with information relevant to the design of BU-manager incentive compensation plans.

The paper proceeds as follows. The next section sets forth the hypotheses tested in this study. Section III outlines the empirical models and estimation procedures used in testing these hypotheses. Section IV provides a discussion of the data and sample, including a brief overview of the survey used to obtain proprietary compensation data. Results are presented in Section V. The final section provides the conclusion and a discussion of the study's limitations.

II. HYPOTHESIS DEVELOPMENT

Newman (1989), Carnes and Guffey (2000), and Atwood et al. (1998) investigate firms' choices of after-tax earnings as the contracting variable in CEO bonus plans. These studies hypothesize that firms with greater tax-planning opportunities, consistent with the Antle and Demski (1988) controllability principle, are more likely to use after-tax performance measures. Using proxies for tax-planning opportunities, these studies collectively find that multinational status, number of operating segments, firm size, and capital intensity are positively associated with after-tax CEO compensation. Atwood et al. (1998) also presents evidence that leverage is negatively associated with this choice.

Dhaliwal et al. (2000) also focuses on the relation between a firm's tax-planning opportunities and the decision whether to use after-tax earnings in CEO bonus plans, but measures planning opportunities using only realized tax credits and absolute values of firms' permanent book-tax differences. They exclude temporary book-tax differences such as those resulting from accelerated tax depreciation because these differences do not impact a firm's after-tax accounting earnings, the performance measure used in after-tax accounting-based bonus plans. Similarly, the ETR used in the current study is computed using total tax expense in the numerator, thereby excluding the effects of temporary book-tax differences.

Consistent with prior literature, Dhaliwal et al. (2000) find that firms with greater tax credits and absolute values of permanent book-tax differences are more likely to compensate their CEOs on an after-tax basis. Because permanent differences can either increase or decrease taxable income relative to book income, they provide no direct evidence of an association between after-tax CEO compensation and lower ETRs. The current study investigates this relation between after-tax CEO compensation and ETRs. The first hypothesis, stated in alternative form, is:

> **H1:** Using after-tax performance measures in CEO accounting-based bonus plans leads to lower ETRs.

This paper also extends the pre- versus after-tax compensation analysis to BU-manager incentive compensation plans. The need for BU-manager involvement in tax-planning efforts is consistent with the coordination task that arises when specialists require input and cooperation from nonspecialists (Milgrom and Roberts 1992). The firm's tax professionals not only need help from the firm's BU managers in the identification of tax-planning opportunities, but also their cooperation in developing and executing strategies to capitalize on such opportunities.

In the context of lowering the ETR and increasing after-tax accounting income, how do BU managers participate in efforts that increase tax credits and lead to permanent differences that lower taxable income? Colgate-Palmolive's CFO states that "Colgate managers routinely watch for ways to time remittances from overseas [which leads to higher foreign tax credits], maximize research and development tax credits...." (Mintz 1999, 2.) BU managers can also help tax professionals formulate transfer prices (Ernst & Young 1999), which helps lower the ETR by decreasing (increasing) income reported in countries

1997.[9] *CAPINT* is computed as the ratio of property, plant, and equipment to the book value of total assets (*BVA*) as of year-end 1997 and is included to control for opportunities related to investments in fixed assets. Larger firms, through the scope and scale of their operations, are expected to have greater opportunities to engage in tax-planning efforts. *SIZE* is computed as the natural logarithm of the *BVA* as of year-end 1997. Firms with greater leverage are also expected to have greater opportunities to engage in tax-planning efforts related to financing activities (Atwood et al. 1998). *LEV* is the ratio of long-term debt to *BVA* as of year-end 1997. Finally, tax laws vary on an industry-by-industry basis. *IND* is a vector of industry indicator variables for the mining and construction, transportation and communication, utilities, sales, financial institutions and insurance, and services industries.

Other explanatory variables are also included in all three equations. *ROA*, the ratio of the average pre-tax income for the three years ending in 1997 to *BVA* as of year-end 1997, is included in the *ETR* equation because prior research (e.g., Wilkie 1988; Gupta and Newberry 1997) has linked ETRs to profitability. *ROA* is also included in the *CEOATAX* and *BUATAX* equations consistent with the Dhaliwal et al. (2000) suggestion that more profitable firms are more likely to compensate the CEO on an after-tax basis because such firms have more income subject to tax. Dhaliwal et al. (2000) also suggests that firms with greater variability in profitability have a greater risk of making business decisions that are inappropriate for their tax status and thus have greater incentives to engage in tax planning. *VARROA*, the variance in *ROA* for the three-year period ended in 1997, is included in both compensation choice equations. *VARROA* is also included in the *ETR* equation because firms with greater incentives to engage in tax planning are expected to have lower ETRs. Bankman (1994) finds that high-growth firms typically place less emphasis on tax-planning efforts, suggesting that such firms are expected to have higher ETRs and are less likely to use after-tax performance measures. *GROWTH*, calculated as the geometric mean annual growth rate of the market value of assets (*MVA*) for the three years 1995–1997, is included in the *ETR* equation and both pre- versus after-tax performance measure choice equations.[10]

Finally, all three equations contain proxies for stock-based compensation. Consistent with the argument that firms with little or no tax liability because of their ESO deductions have little incentive to use after-tax performance measures, Dhaliwal et al. (2000) find that firms with higher proportions of current CEO compensation comprised of stock option grants are less likely to use after-tax earnings in CEO incentive compensation plans. Proxies for stock-based compensation are thus included in both pre- versus after-tax performance measure choice equations. *CEOESO*, the percentage of the CEO's compensation received in 1997 attributable to stock option grants calculated as the ratio of the Black-Scholes value (BSV) of stock option grants to the sum of the CEO's 1997 salary, annual bonus, and BSV of stock option grants, is included in the *CEOATAX* equation.[11] Similarly, *BUSTK*, an indicator variable equal to 1 if stock price is used as a performance measure in a majority of the firm's business-unit managers' incentive compensation plans, is included in the

---

[9] Under SFAS No. 14, which was in effect in 1997, geographic segment disclosures are required only if a firm's multinational operations meet a 10 percent materiality threshold. To the extent a firm responded that it filed one or more foreign tax returns but reported no foreign assets in its 1997 10-K, a 5 percent ratio of foreign-to-total assets is assumed.

[10] *MVA* is computed as the book value of assets minus the book value of equity plus the product of year-end stock price times the number of shares of common stock outstanding at year-end.

[11] 1997 salary, annual bonus, and BSV of options granted were obtained from firms' proxy statements. When the BSV of options granted was not disclosed, the BSV of options granted was computed as the number of options granted to the CEO multiplied by the weighted average per share value of options granted to all employees as in the disclosed FAS No. 123 footnote contained in the firms' annual reports.

BUATAX equation. Both *CEOESO* and *BUSTK* are included in the *ETR* equation to control for the possibility that stock-based compensation motivates CEOs and BU managers to focus on after-tax earnings and engage in tax-planning efforts that result in lower ETRs.[12]

Exogenous Variables Appearing In Only One Equation

Following the Mills et al. (1998) finding that ETRs are negatively related to spending on the tax function, *LSPEND* is included in the *ETR* equation. *LSPEND* is computed as the natural logarithm of 1 plus total 1997 spending on the corporate tax function scaled by total SGA expense. *LSPEND* is not included in the two compensation choice equations because a firm's decisions to compensate its CEO and BU managers on an after-tax basis are not decisions typically made on a year-by-year basis and are thus not expected to be based on the tax department's annual budget.[13]

Bushman et al. (1995) and Keating (1997) find that the weight placed on firm-wide performance measures increases as interdependencies among a firm's BUs increase. Such interdependencies could lead to a greater need for a firm's BU managers to coordinate tax-planning efforts, such as those relating to transfer pricing. Accordingly, when a firm uses a firm-wide, and not a unit-based, accounting-based performance measure in its BU-manager incentive compensation plans, this performance measure is more likely to be computed on an after-tax basis. *FWONLY* is defined as an indicator variable equal to 1 if the firm uses a firm-wide accounting-based performance measure, but not a BU-based measure, in its BU-manager incentive compensation plans, and 0 otherwise. *FWONLY* is included in the *BUATAX*, but not the *CEOATAX* equation, because the need to coordinate efforts among a firm's BU managers does not apply to firm-level decisions.

Specification

Several specification issues arise in connection with this three-equation double treatment effects model. First, the ETR in the absence of compensating managers on an after-tax basis (hereinafter ETR*) does not appear in either the *CEOATAX* or *BUATAX* selection model. A higher ETR* neither impacts the potential benefit of ETR-lowering efforts nor enhances the opportunities for managers to engage in tax-planning efforts. Instead, following the Antle and Demski (1988) controllability principle, it is the manager's ability to take actions that lower the ETR that is critical to the decision to use pre- versus after-tax compensation. Arguably, this ability is unrelated to a firm's ETR in the absence of using after-tax performance measures.

Second, although these three equations are not modeled as a three-equation simultaneous-equation system, assignment of firms to subsamples based on whether firms use after-tax performance measures is not random. As a result, the *CEOATAX* and *BUATAX* variables are potentially endogenous to the *ETR* because the expected *ETR* equation error term given the use of after-tax compensation may not equal zero, e.g., if any tax-planning opportunities that influence both the ETR and the pre- versus after-tax compensation decisions are omitted from the three equations. Accordingly, the *ETR* equation includes the two selectivity correction variables that help correct for the possible bias in all *ETR* model coefficient estimates resulting from this potential endogeneity. In the *ETR* equation, the coefficients on *CEOATAX* and *BUATAX* represent these variables' incremental effect on the ETR after controlling for their endogeneity, i.e., after removing the effect of the exogenous determinants of *CEOATAX*, *BUATAX*, and the *ETR*.

---

[12] *CEOESO* and *BUSTK*, along with *LSPEND*, could be modeled as a function of *CEOATAX* and as endogenous to the *ETR* equation. Arguably, however, modeling the *CEOATAX* and *BUATAX* variables as endogenous helps correct for much of the potential endogeneity bias associated with the explanatory variables included in the *ETR* equation.

[13] As a sensitivity test, Equations (2) and (3) are reestimated including *LSPEND* as an explanatory variable in both equations. *LSPEND* is insignificant in both estimations.

A third specification issue relates to the potentially sequential nature of a firm's decisions to compensate its CEO and BU managers on a pre- versus after-tax basis. Arguably, a firm's directors first decide whether to compensate its CEO using after-tax earnings. The CEO then decides whether to base BU manager compensation on after-tax results. The data in Table 2 show that firms typically do not compensate their BU managers on an after-tax basis unless the CEO is also compensated after-tax, suggesting that these two decisions are sequential in nature. *CEOATAX* is thus included as an explanatory variable in the *BUATAX* equation, but not *vice versa*. Taken together, the *CEOATAX* and *BUATAX* equations comprise a recursive, simultaneous binary choice model (Greene 1998a; 2003, 215).[14]

The final specification issue relates to the identification of the coefficients and covariance parameters in Equations (1)–(3). Based on the discussion in Tunali (1986, 245), the *CEOATAX* and *BUATAX* selection models, Equations (2) and (3), are identified because all four possible combinations of the two choice variables (after-tax compensation for both CEO and BU managers, after-tax for CEO and pre-tax for BU managers, pre-tax for CEO and after-tax for BU managers, and pre-tax for both CEO and BU managers) are observed. This complete 2 × 2 classification results in 3 degrees of freedom, which allows the two means and correlation coefficient to be identified. The *ETR* model parameters are identified because only this model includes *LSPEND*, the measure of tax function spending, and the selectivity correction variables, λ_{CEO} and λ_{BU}, which are nonlinear functions of the unknown parameters in the *CEOATAX* and *BUATAX* selection models.

Estimation

In general, OLS coefficient estimates will be inconsistent if the error terms of the outcome (treatment) equation and the selection equation are correlated (Heckman 1976). Tunali (1986) formulates a general double-selection model in which the typical Heckman (1976) selection model is extended by including a second selection equation. This paper follows Tunali's (1986) two-step approach in which the *CEOATAX* and *BUATAX* selection equations are first estimated using maximum likelihood bivariate probit.

Estimation of the *ETR*, *CEOATAX*, and *BUATAX* models proceeds as follows. First, this three-equation double treatment effects model is expressed as:

$$ETR = \beta' X_1 + \mu, \tag{1}$$

$$CEOATAX^* = \gamma' X_2 + \varepsilon, \tag{2}$$

$$BUATAX^* = \alpha' X_3 + \varphi, \tag{3}$$

where X_1, X_2, and X_3 are the vectors of explanatory variables in the respective structural equations, β', γ', and α' are the vectors of coefficients, and μ, ε, and φ, are error terms that are assumed to have a trivariate normal distribution with variances equal to σ^2, 1, and 1, respectively, and correlations ρ_{12}, ρ_{13}, and ρ_{23}. *CEOATAX** and *BUATAX** are unobserved latent variables. Instead, the following variables are observed:

---

[14] Following Greene (1998a; 2003, 383), the *CEOATAX* and *BUATAX* equations are recursive because after-tax CEO compensation is determined solely by exogenous variables. Then, given *CEOATAX*, after-tax BU-manager compensation is also a function of exogenous variables. The two equations are jointly determined via the direct effect of including *CEOATAX* in the *BUATAX* equation and indirectly through the equations' error terms.

$$CEOATAX = 1 \text{ if } CEOATAX^* > 0, 0 \text{ otherwise; and}$$
$$BUATAX = 1 \text{ if } BUATAX^* > 0, 0 \text{ otherwise.}$$

Note that *CEOATAX* is included in X_3, the vector of explanatory variables in the *BUATAX* equation, and X_1, the vector of explanatory variables in the *ETR* equation, includes both *CEOATAX* and *BUATAX*.

Bivariate probit jointly estimates the *CEOATAX* and *BUATAX* models utilizing the correlation in error terms, ρ_{23}, and produces the selectivity correction variables, λ_{CEO} and λ_{BU}, which represent the double-selection model analogs to the Heckman (1976) single-equation inverse Mill's ratio.[15] The selectivity correction variables, λ_{CEO} and λ_{BU}, are calculated as:

$$\lambda_{CEO} = \phi(-\gamma' X_2)\Phi[(-\alpha' X_3 - \rho_{12} CEOATAX)/(1 - \rho_{12}{}^2)^{1/2}]/\Phi_2, \text{ and}$$
$$\lambda_{BU} = \phi(-\alpha' X_3)\Phi[(-\gamma' X_2 - \rho_{13} BUATAX)/(1 - \rho_{13}{}^2)^{1/2}]/\Phi_2$$

where:

$\phi(\bullet)$ = PDF;
$\Phi(\bullet)$ = CDF; and
Φ_2 = bivariate normal CDF, $\Phi(-\gamma' X_2, -\alpha' X_3, \rho_{23})$.

Finally, OLS regression of the *ETR* on *CEOATAX*, *BUATAX*, other explanatory variables, λ_{CEO}, and λ_{BU} produces consistent coefficient estimates.[16]

The above estimation approach is a restricted control function (CF) approach because the outcome equation includes the selectivity correction variables and the treatment effect is captured only in a shift in the outcome equation's intercept (e.g., Vella and Verbeek 1999). In contrast to the restricted CF approach, an instrumental variables approach imposes the restrictions that either the treatment effect does not vary cross-sectionally or that a firm's idiosyncratic benefit is not known *ex ante* (Heckman 1997), neither of which is likely satisfied in the current setting. Vella and Verbeek (1999) find that the CF approach is more efficient and robust to specification errors, but that it is more dependent on the normality assumption. They argue that the restricted CF approach is more appropriate in settings such as the current one in which a treatment's associated costs are excluded from the choice model.[17]

IV. SAMPLE AND DATA

Proprietary survey data are used to construct *CEOTAX* and *BUATAX*, the variables indicating whether the firm uses after-tax performance measures in its CEO and BU-manager incentive compensation plans, respectively. Survey data are also used to construct *BUSTK*, the indicator of BU-manager stock-based compensation, *LSPEND*, the amount spent on the corporate tax function, and to determine whether a firm has multiple BUs. The

---

[15] If the *CEOATAX* equation were the only selection equation, the selectivity correction variable, λ_{CEO}, would be calculated as $\phi(-\gamma' \mathbf{X}_2)/\Phi(-\gamma' \mathbf{X}_2)$.

[16] Estimation of the augmented Equation (1) using *LIMDEP* produces corrected standard errors. See Greene (1998b, 740) for details on how the asymptotic covariance matrix is computed.

[17] One limitation of the two-step estimation approach is that it could lead to a loss of efficiency relative to using maximum likelihood procedures (Wales and Woodland 1980). In a single-selection setting, Nelson (1984) shows that efficiency loss is greatest when the exogenous variables in the selection and outcome equations are highly correlated. Because most of the exogenous *ETR* equation explanatory variables are also included in the *CEOATAX* and *BUATAX* selection equations, efficiency loss could result.

Appendix contains the survey's general instructions and survey questions relevant to these variables. All other variables, including the *ETR*, are constructed using Compustat data.

Questionnaires were sent to corporate executives selected from two sets of firms that met preliminary data eligibility requirements. The first set was drawn from a list provided by the Financial Executives Institute. The second set was drawn from *Fortune* 500 firms not included in the first set and for which a tax executive's name and address was listed in the *Tax Analysts* corporate tax manager directory. Tax executives were targeted because they were presumably more interested in this study's results and were thus more likely to respond to the questionnaire (Dillman 1978). Survey instruments were sent to 829 publicly traded firms that met preliminary Compustat data availability requirements and had positive pre-tax income before equity in subsidiaries earnings in both 1995 and 1996.[18] All potential respondents were sent postcard reminders three weeks after the initial mailing. Thereafter, second and third reminders along with the questionnaire and a postage-paid return envelope were sent to nonrespondents. This process yielded 249 survey responses, resulting in a 30 percent response rate. Nine responses from firms that did not meet final data requirements and three unidentified responses were excluded, leaving 237 usable instruments.

To investigate how the 829 firms surveyed (survey firms) compare to the remaining population of firms listed in Compustat (non-survey firms), descriptive statistics comparing the two groups were computed. These comparative statistics (not tabulated) indicate that survey firms are larger, more capital intensive, more leveraged, and have higher ETRs than non-survey firms.[19] Targeting larger firms with a tax executive was necessary to produce a response rate adequate to test this study's hypotheses. Accordingly, the results may not be generalizable to smaller firms. To the extent, however, that small firms are less likely to have multiple business units and more likely to have net operating losses (Wang 1991), and thus less need for tax-planning efforts, the inability to generalize the results to small firms may not be an issue.

To investigate whether nonresponse bias is an issue, descriptive statistics were computed to compare firms responding to the survey (respondent firms) and those firms that did not respond (nonrespondent firms). This comparison (not tabulated) reveals no statistically significant mean differences in the *ETR*, pre-tax profitability (*ROA*), size (*BVA*), leverage (*LEV*), capital intensity (*CAPINT*), and the market-to-book ratio, suggesting that nonresponse bias is not a concern.

Table 1, Panel A summarizes the sample selection process that results in a final sample of 209 firms. Fourteen firms without multiple BUs were excluded. Eight firms that did not base the CEO's bonus on an accounting measure were also excluded along with five firms that used neither a unit-based nor a firm-wide accounting measure to compensate its BUs managers. One outlier with an *ETR* in excess of 300 percent was also deleted.[20] Table 1, Panel B summarizes the final sample's industry composition.

---

[18] 1997 Compustat data were not available at the time survey firms were selected.

[19] Survey (non-survey) firms have an *ETR* of .353 (.314), *BVA* of 7,769 million (1,149 million), *LEV* of .179 (.160), and *CAPINT* of .311 (.227). All mean differences are significant at the .05 level in a two-tailed test.

[20] The income tax footnotes of firms with extreme ETR observations were reviewed to identify the sources of extreme ETRs. Extreme ETRs were recomputed to remove the effects of large, nonrecurring statutory reconciliation items such as the effects of a business dispositions and asset impairments, and decreases in the deferred tax asset valuation allowance. The results based on estimating the model without these eight observations with extreme ETR observations prior to correction (n = 201) are substantially the same as those reported in Tables 4 through 6.

TABLE 1
Sample Determination and Industry Classification

Panel A: Sample Determination

Number of firms responding to survey	249
Less: Responses that did not identify the firm	3
Less: Observations that do not have data available on either the 1997 Compustat or SEC EDGAR database	9
Usable Instruments	237
Less: Firms without multiple business units	14
Less: Firms not basing the CEO's annual bonus on accounting results	8
Less: Firms not using either a firm-wide or BU accounting measure in the incentive compensation plans of its BU managers	5
Less: Outlier	1
Final Sample	209

Panel B: Industry Classification

Industry	SIC Codes Ranging From:	
Mining and Construction	1000 to 1999	6
Manufacturing	2000 to 3999	114
Transportation and Communication	4000 to 4899	9
Utilities	4900 to 4999	20
Sales	5000 to 5999	24
Financial Institutions and Insurance	6000 to 6699	24
Services	7000 to 9099	12
Final Sample		209

V. RESULTS

Descriptive Statistics

Table 2, Panel A presents descriptive statistics for sample firms. The mean *ETR* (.356) approximates the top statutory corporate tax rate. The means for *CEOATAX* and *BUATAX*, respectively, indicate that approximately 61 percent of the sample firms compensate the CEO on an after-tax basis while approximately 32 percent use an after-tax performance measure to compensate their BU managers.[21]

Table 2, Panel B presents means for the two subsamples based on whether the CEO is compensated on a pre-tax (n = 81) versus after-tax (n = 128) basis. The *ETR* mean when *CEOATAX* equals 1 (.348) is significantly lower than the mean when *CEOATAX* equals 0 (.368), consistent with H1. Consistent with prior research, the means for *BVA* (firm size) and *CAPINT* (capital intensity) are greater for the after-tax firms. The *BUATAX* mean difference indicate that few sample firms compensate BU managers on an after-tax basis when the CEO is compensated using a pre-tax measure, and this percentage increases substantially when the CEO's bonus is based on an after-tax measure. This result provides evidence that a firm is more likely to compensate its BU managers on an after-tax basis if the CEO is compensated after-tax, consistent with these decisions being sequential in nature. Contrary to Dhaliwal (2000), the mean percentage of CEO compensation derived from stock option grants is higher, not lower, for the after-tax firms.

---

[21] Seventy percent of the sample firms in Atwood et al. (1998) use after-tax earnings in the CEO's incentive compensation plan. Samples in other studies (e.g., Newman 1989; Dhaliwal et al. 1999) have a considerably lower percentage of after-tax firms. Atwood et al. (1998), unlike these other studies, code a bonus plan as after-tax when the firm's proxy statement uses "net income" or related terminology that is silent regarding the pre-tax versus after-tax nature of this amount. They assume terms such as "net income" means after-tax income unless otherwise stated.

TABLE 2
Descriptive Statistics

Panel A: Overall Descriptive Statistics

Continuous Variables	Mean	Std. Dev.	10%	25%	Median	75%	90%
ETR	.356	.062	.279	.325	.360	.385	.416
SPEND	.006	.011	.001	.001	.003	.006	.011
MULTI	.171	.185	.000	.050	.072	.278	.429
CAPINT	.322	.225	.017	.152	.285	.485	.653
LEV	.178	.133	.006	.066	.162	.275	.344
BVA	7,371	12,962	265	730	2,029	6,694	21,453
ROA	.107	.070	.024	.054	.092	.148	.196
VARROA	.001	.002	.000	.000	.001	.002	.003
GROWTH	1.213	.223	.999	1.082	1.173	1.293	1.498
CEOESO	.315	.243	.000	.086	.305	.468	.671

Dichotomous Variables	Mean	Std. Dev.
CEOATAX	.612	.488
BUATAX	.316	.466
FWONLY	.120	.325
BUSTK	.603	.490

Panel B: Subsample Means—Pre-Tax versus After-Tax CEO Annual Bonus

Variable	Mean for Subsample in which *CEOATAX* = 0 (n = 81)	Mean for Subsample in which *CEOATAX* = 1 (n = 128)	t-test of Differences in Means
ETR	.368	.348	2.268*
BUATAX	.111	.445	5.927*
SPEND	.006	.014	0.945
MULTI	.151	.184	0.200
CAPINT	.281	.348	2.147*
LEV	.177	.179	0.126
BVA	4,817	9,013	2.541*
ROA	.103	.109	0.637
VARROA	.001	.001	0.517
GROWTH	1.216	1.211	0.138
CEOESO	.267	.345	2.387*

(continued on next page)

TABLE 2 (continued)

Panel C: Subsample Means—Pre-Tax versus After-Tax BU-Manager Compensation

Variable	Mean for Subsample in which *BUATAX* = 0 (n = 143)	Mean for Subsample in which *BUATAX* = 1 (n = 66)	t-test of Differences in Means
ETR	.363	.340	2.494*
CEOATAX	.497	.864	6.143*
SPEND	.013	.005	1.051
MULTI	.160	.195	1.195
CAPINT	.298	.375	2.233*
LEV	.178	.177	0.070
BVA	5,779	10,848	2.351*
ROA	.111	.096	1.543
VARROA	.001	.001	0.517
GROWTH	1.216	1.211	.138
FWONLY	.056	.258	3.503*
BUSTK	.559	.697	1.948

---

* Significant at the .05 level (two-tailed test).

Variable Definitions:

ETR = average annual ETR, defined as the ratio of total tax expense to pre-tax income, for the three years ending in 1997;
CEOATAX = 1 if the CEO's annual bonus is based on after-tax accounting results;
BUATAX = 1 if the firm's BU managers are compensated on an after-tax basis;
SPEND = 1997 spending on the corporate tax function scaled by total SGA expense;
BVA = book value of total assets as of year-end 1997 (000s omitted);
MULTI = the ratio of foreign assets to *BVA* as of year-end 1997;
CAPINT = the ratio of property, plant, and equipment to *BVA* as of year-end 1997;
LEV = the ratio of long-term debt to *BVA* as of year-end 1997;
ROA = the ratio of the average pre-tax income for the three years ending in 1997 to *BVA* as of year-end 1997;
VARROA = the variance in ROA for the three-year period ended in 1997;
GROWTH = the geometric mean annual growth rate of the market value of assets (*MVA*) for the three years 1995–1997;
FWONLY = 1 if the firm uses a firm-wide accounting measure, but not a unit-based measure, in its BU-manager incentive compensation plans;
CEOESO = the percentage of the CEO's compensation received in 1997 attributable to stock option grants calculated as the ratio of the value of stock option grants to the sum of the CEO's salary, annual bonus, and value of stock option grants; and
BUSTK = 1 if stock price is used as a performance measure in the incentive compensation plans of a majority of the firm's business-unit managers.

Table 2, Panel C presents means for the two subsamples based on whether the firm's BU managers are compensated on a pre-tax (n = 143) versus after-tax (n = 66) basis. Consistent with H2, the mean ETR for the after-tax firms is significantly lower than the mean ETR for the pre-tax firms. Firms using after-tax BU performance measures are larger and more capital-intensive than pre-tax firms. Consistent with expectations, the mean difference in *FWONLY* indicates that firms using a firm-wide accounting measure, and not a unit-based measure, to compensate its BU managers are more likely to base this measure on after-tax results.

Simple Correlations

Table 3 contains simple correlations among the test variables with significance levels noted at the .05 level. Consistent with H1 and H2, the *ETR* is significantly negatively correlated with *CEOATAX* and *BUTATAX*. The only other significant correlations involving the *ETR* are those with *LSPEND* (spending on the tax function), *LEV* (leverage), and *FWONLY*. The correlations involving *CEOATAX* and *BUATAX* are consistent with the subsample mean differences reported in Table 2, Panels B and C.

ETR Model Estimations

Table 4 presents the *ETR* model estimation results. OLS results, presented in Column (1), do not support either H1 or H2. Without correcting for potential endogeneity bias, the coefficients on *CEOATAX* (β_1) and *BUATAX* (β_2) are both negative but insignificant.

Table 4, Column (2), reflects the results of estimating the *ETR* model augmented with the selectivity correction variables, λ_{CEO} and λ_{BU}, which are calculated based on the results of the bivariate probit estimation of the *CEOATAX* and *BUATAX* models. The coefficient on *CEOATAX* (β_1) is insignificant and thus fails to support H1. The coefficient on *BUATAX* (β_2), however, is −.071 and significant ($p = .005$). This result is consistent with H2 and provides evidence of a link between after-tax BU-manager performance measures and lower ETRs.[22]

The economic significance of the association between after-tax BU-manager compensation and lower ETRs is based on the interpretation of the coefficient on *BUATAX*, β_2, as the "experimental average treatment effect" (Heckman 1990). Although firms are not randomly assigned this treatment, the estimated intercept shift, β_2, captures the average effect on the ETR of a random assignment of the after-tax BU-manager compensation treatment. Using the estimated coefficient on *BUATAX*, −.071, per-firm tax savings are computed, resulting in a mean (median) tax savings of $41.9 million ($13.3 million), relative to mean (median) total tax expense of $211.2 ($64.8) million.[23] This evidence is consistent with the argument that compensating BU managers on an after-tax basis is not only statistically associated with lower ETRs, but also results in economically significant tax savings.

The result that after-tax compensation for BU managers, but not CEOs, is associated with lower ETRs leads to at least two questions. First, given the economic significance of the tax savings that result from compensating BU managers on an after-tax basis, why do a majority of firms compensate their BU managers on a pre-tax basis? Following the Antle and Demski (1988) controllability principle, the most plausible explanation for this behavior is that for firms compensating their BU managers on a pre-tax basis the expected costs associated with after-tax compensation outweigh the expected benefits. Second, why is CEO after-tax compensation not associated with lower ETRs? It is likely that CEOs have other incentives that sufficiently motivate a focus on after-tax, rather than pre-tax, earnings. For example, a CEO's job security could depend on after-tax earnings if the firm's shareholders expect the CEO to maintain a certain profitability level. To the extent, however, that CEOs compensated on an after-tax basis are more likely to implement after-tax BU-manager

---

[22] As previously discussed, the ETR does not reflect the tax benefits of ESO deductions. To the extent that firms with large ESO deductions are less likely to use after-tax compensation, adjusting the ETRs to reflect these tax benefits would bias toward failing to reject the null hypotheses of no association between ETRs and after-tax performance measures.

[23] A firm's tax savings from using after-tax BU performance measures was computed as the product of the coefficient on β_2, −.071, and that firm's pre-tax income. Mean (median) per-firm tax savings reported above represent the mean (median) tax savings for the 209 sample firms.

TABLE 3
Correlations

	ETR	*BUATAX*	*CEOATAX*	*LSPEND*	*MULTI*	*CAPINT*	*LEV*	*SIZE*	*ROA*	*VARROA*	*GROWTH*	*FWONLY*	*CEOESO*
ETR	1.000												
BUATAX	−.169*	1.000											
CEOATAX	−.157*	.350*	1.000										
LSPEND	−.158*	−.004	−.010	1.000									
MULTI	.053	.088	.087	−.061	1.000								
CAPINT	.042	.160*	.146*	−.021	−.058	1.000							
LEV	.147*	−.005	.009	−.048	−.076	.387*	1.000						
SIZE	−.070	.197*	.162*	−.130	.130	.008	.221*	1.000					
ROA	.053	−.102	.046	−.047	.255*	.045	−.288*	−.289*	1.000				
VARROA	−.070	−.019	−.035	−.016	.048	−.074	−.202*	−.304*	.246*	1.000			
GROWTH	−.029	−.087	−.010	.126	.081	−.124	−.118	−.043	.293*	.160*	1.000*		
FWONLY	−.163*	.289*	.051	−.031	−.060	.112	−.047	−.124	.127	.111	.079	1.000	
CEOESO	−.040	.132	.158*	−.130	.165*	−.063	−.021	.317*	.090	.003	.081	.092	1.000
BUSTK	.044	.131	.137*	−.156*	.030	.034	.119	.287*	.056	−.016	−.061	.028	.111

* Significant at the .05 level.

Variable Definitions:

ETR = average annual ETR, defined as the ratio of total tax expense to pre-tax income, for the three years ending in 1997;
BUATAX = 1 if the firm's BU managers are compensated on an after-tax basis;
CEOATAX = 1 if the CEO's annual bonus is based on after-tax accounting results;
LSPEND = the natural logarithm of 1 plus total 1997 spending on the tax function scaled by total SGA expense;
MULTI = the ratio of foreign assets to the book value of total assets (*BVA*) as of year-end 1997;
CAPINT = the ratio of property, plant, and equipment to *BVA* as of year-end 1997;
LEV = the ratio of long-term debt to *BVA* as of year-end 1997;
SIZE = the natural logarithm of *BVA* as of year-end 1997;
ROA = the ratio of the average pre-tax income for the three years ending in 1997 to *BVA* as of year-end 1997;
VARROA = the variance in *ROA* for the three-year period ended in 1997;
GROWTH = geometric growth rate of the market value of assets for the three years 1995–1997;
FWONLY = 1 if the firm uses a firm-wide accounting measure, but not a unit-based measure, in its BU-manager compensation plans;
CEOESO = the percentage of the CEO's compensation received in 1997 attributable to stock option grants calculated as the ratio of the value of stock option grants to the sum of the CEO's salary, annual bonus, and value of stock option grants; and
BUSTK = 1 if stock price is used as a performance measure in the incentive compensation plans of a majority of the firm's business-unit managers.

TABLE 4
ETR Model Estimations

$$ETR_i = \beta_0 + \beta_1 CEOATAX_i + \beta_2 BUATAX_i + \beta_3 LSPEND_i + \beta_4 MULTI_i + \beta_5 CAPINT_i + \beta_6 LEV_i + \beta_7 SIZE_i + \beta_8 ROA_i + \beta_9 VARROA_i + \beta_{10} GROWTH_i + \beta_{11} CEOESO_i + \beta_{12} BUSTK_i + \beta_{13-18} IND_i + \theta_{CEO} \lambda CEO_i + \theta_{BU} \lambda BU_i + \mu_i$$

			(1)[a]		(2)[b]	
Independent Variable	**Coeff.**	**Pred. Sign**	**Coeff. Est.**	**p-value[c]**	**Coeff. Est.**	**p-value[c]**
Intercept	β_0	?	0.391	.000	0.386	.000
CEOATAX	β_1	−	−0.010	.152	0.063	.660
BUATAX	β_2	−	−0.011	.129	−0.071	.005
LSPEND	β_3	−	−0.617	.067	−0.675	.048
MULTI	β_4	?	−0.022	.393	−0.017	.614
CAPINT	β_5	?	−0.031	.247	−0.032	.667
LEV	β_6	?	0.028	.491	0.007	.863
SIZE	β_7	?	−0.002	.471	−0.003	.583
ROA	β_8	+	−0.012	.559	−0.110	.774
VARROA	β_9	−	−5.708	.010	−5.519	.036
GROWTH	β_{10}	+	0.005	.398	0.004	.429
CEOESO	β_{11}	−	0.009	.838	0.000	.504
BUSTK	β_{12}	−	0.009	.915	0.013	.920
λ_{CEO}	θ_{CEO}	?	NA		−0.035	.708
λ_{BU}	θ_{BU}	?	NA		0.044	.053
Model Chi-Squared			2.34 (190 d.f.)		2.43 (188 d.f.)	
Adj./Pseudo R^2			.104		.121	

[a] Column (1) presents the results from estimating the ETR model using OLS with all variables assumed exogenous and excluding λ_{CEO} and λ_{BU} from the model.

[b] Column (2) presents the results from estimating the ETR model augmented with the selectivity correction (λ) variables from the *CEOATAX* and *BUATAX* selection equations. The correlation between the error terms of the ETR equation and the *CEOATAX* (*BUATAX*) selection equation is −.495 (.638).

[c] p-values are based on a one- (two-) tailed test where the coefficient sign is (not) predicted.

Variable Definitions:

ETR = average annual ETR, defined as the ratio of total tax expense to pre-tax income, for the three years ending in 1997;
CEOATAX = 1 if the CEO's annual bonus is based on after-tax accounting results;
BUATAX = 1 if the firm's BU managers are compensated on an after-tax basis;
LSPEND = the natural logarithm of 1 plus total 1997 spending on the tax function scaled by total SGA expense;
MULTI = the ratio of foreign assets to the book value of total assets (*BVA*) as of year-end 1997;
CAPINT = the ratio of property, plant, and equipment to *BVA* as of year-end 1997;
LEV = the ratio of long-term debt to *BVA* as of year-end 1997;
SIZE = the natural logarithm of *BVA* as of year-end 1997;
ROA = the ratio of the average pre-tax income for the three years ending in 1997 to *BVA* as of year-end 1997;
VARROA = the variance in *ROA* for the three-year period ended in 1997;
GROWTH = the geometric mean annual growth rate of the market value of assets (*MVA*) for the three years 1995–1997;
CEOESO = the percentage of the CEO's compensation received in 1997 attributable to stock option grants calculated as the ratio of the value of stock option grants to the sum of the CEO's salary, annual bonus, and value of stock option grants;
BUSTK = 1 if stock price is used as a performance measure in the incentive compensation plans of a majority of the firm's business-unit managers;
IND = vector of industry indicator variables for the mining and construction, transportation and communication, utilities, sales, financial institutions and insurance, and services industries (results not tabulated); and
λ_{CEO} (λ_{BU}) = the selectivity correction variable from the *CEOATAX* (*BUATAX*) equation resulting from the bivariate probit estimation of the *CEOATAX* and *BUATAX* equations.

performance measures that are directly linked to lower ETRs, the evidence is consistent with an indirect link between after-tax CEO compensation and lower ETRs.

Other *ETR* model estimation results reported in Table 4 reveal that the relation between the *ETR* and firm spending on the tax function (*LSPEND*) is significantly negative, consistent with Mills et al. (1998). The coefficient on the variation in *ROA* (*VARROA*) is also significantly negative. In contrast to previous studies (e.g., Gupta and Newberry 1997; Mills et al. 1998) the coefficients on other explanatory variables are insignificant. There are at least three possible explanations for these results. First, prior ETR studies typically investigate determinants using the ETR specification that has only current tax expense in the numerator, whereas the ETR in this study is computed as the ratio of total tax expense to pre-tax income. Explanatory variables used in prior studies could possibly explain only variation in current tax expense; e.g., the capital intensity measure, *CAPINT*, could proxy solely for temporary depreciation book-tax differences. Second, the coefficients on these explanatory variables in the *ETR* model estimation only capture their incremental direct effects above and beyond their indirect effect on the *ETR* via their inclusion in the *CEOATAX* and *BUATAX* selection equations (Greene 2003, 783). Accordingly, the insignificance of the ETR determinants could be due to the separation of these direct and indirect effects. Finally, as reported in Table 3, the significant correlations between *MULTI*, *CAPINT*, *LEV*, *SIZE*, and *ROA* could lead to higher standard errors and lower t-statistics associated with these variables' coefficient estimates.

The coefficients on the selectivity correction variables, λ_{CEO} and λ_{BU}, are $-.035$ (p $= .758$) and .044 (p $= .053$), respectively. The selectivity correction variables help correct for the possible bias in all *ETR* model coefficient estimates, not just the coefficients on *CEOATAX* and *BUATAX*, resulting from the omission of unobserved variables from the *ETR* model that are correlated with unobserved variables from the *CEOATAX* and *BUATAX* equations. A t-test of these coefficients represents a test of exogeneity (e.g., Vella and Verbeek 1999), suggesting that only *BUATAX* is endogenous to the *ETR* equation.

CEOATAX and *BUATAX* Selection Equation Estimations

Table 5, Column (1) presents the results of estimating the *CEOATAX* model using probit regression. Column (2) of Table 5 presents the results of estimating the *CEOATAX* model in a bivariate probit estimation of the *CEOATAX* and *BUATAX* models. Similarly, Table 6, Columns (1) and (2), present the results of estimating the *BUATAX* model using probit and bivariate probit, respectively. The correlation in error terms between the two selection equations is $-.331$ but insignificant (p $= .813$). This insignificance is due to the inclusion of *CEOATAX* in *BUATAX* equation, consistent with the findings in Greene (1998).

As reported in Table 5 and consistent with prior research, capital intensity (*CAPINT*) is positively associated with the choice to compensate the CEO on an after-tax basis in both the single equation and bivariate probit estimations. The coefficients on firm size (*SIZE*) and profitability (*ROA*) are also positive, though only weakly significant, consistent with previous findings. Contrary to prior research, the coefficients on multinational status (*MULTI*), leverage (*LEV*), and the percentage of the CEO's compensation derived from stock option grants (*CEOESO*) are insignificant.

In the single-equation *BUATAX* model, probit estimation results reported in Table 6, Column (1), the coefficient on *CEOATAX* (.945) is significant (p $= .000$), consistent with the argument that firms compensating their CEOs on an after-tax basis are more likely to use after-tax BU-manager compensation. The coefficient on *CAPINT* (capital intensity) is significantly positive, consistent with expectations. Multinational status (*MULTI*) and firm size (*SIZE*) are also positive, as expected, but their associated coefficients are only weakly

TABLE 5
CEOATAX Model Estimations

$$CEOATAX_i = \gamma_0 + \gamma_1 MULTI_i + \gamma_2 CAPINT_i + \gamma_3 LEV_i + \gamma_4 SIZE_i + \gamma_5 ROA_i + \gamma_6 VARROA_i + \gamma_7 GROWTH_i + \gamma_8 CEOESO_i + \gamma_{9-14} IND_i + \varepsilon_i$$

			(1)[a]		(2)[b]	
Independent Variable	**Coeff.**	**Pred. Sign**	**Coeff. Est.**	**p-value[c]**	**Coeff. Est.**	**p-value[c]**
Intercept	γ_0	?	−1.056	.160	−1.058	.174
MULTI	γ_1	?	0.405	.242	0.430	.247
CAPINT	γ_2	+	1.285	.013	1.247	.019
LEV	γ_3	+	−0.271	.621	−0.210	.588
SIZE	γ_4	+	0.094	.091	0.092	.115
ROA	γ_5	+	2.398	.087	2.360	.098
VARROA	γ_6	+	37.118	.257	37.126	.278
GROWTH	γ_7	−	−0.256	.281	−0.246	.286
CEOESO	γ_8	−	0.638	.936	0.650	.930
Model Chi-Squared (14 d.f.)			25.92		NA	
Pseudo R^2			.093		NA	

[a] Column (1) presents the results from a probit estimation of the *CEOATAX* model.

[b] Column (2) presents results from the maximum likelihood bivariate probit estimation of the *CEOATAX* and *BUATAX* models. The correlation between the error terms of the *CEOATAX* and *BUATAX* selection equations is −.331 (p = .813).

[c] p-values are based on a one- (two-) tailed test where the coefficient sign is (not) predicted.

Variable Definitions:

CEOATAX = 1 if the CEO's annual bonus is based on after-tax accounting results;
MULTI = the ratio of foreign assets to the book value of total assets (*BVA*) as of year-end 1997;
CAPINT = the ratio of property, plant, and equipment to *BVA* as of year-end 1997;
LEV = the ratio of long-term debt to *BVA* as of year-end 1997;
SIZE = the natural logarithm of *BVA* as of year-end 1997;
ROA = the ratio of the average pre-tax income for the three years ending in 1997 to BVA as of year-end 1997;
VARROA = the variance in *ROA* for the three-year period ended in 1997;
GROWTH = the geometric mean annual growth rate of the market value of assets (*MVA*) for the three years 1995–1997;
CEOESO = the percentage of the CEO's compensation received in 1997 attributable to stock option grants calculated as the ratio of the value of stock option grants to the sum of the CEO's salary, annual bonus, and value of stock option grants; and
IND = vector of industry indicator variables for the mining and construction, transportation and communication, utilities, sales, financial institutions and insurance, and services industries (results not tabulated).

significant. The probit results also provide strong evidence that firms using a firm-wide accounting measure (*FWONLY*), and not a unit-based measure, are more likely to compensate BU managers on an after-tax basis.[24]

---

[24] When the *BUATAX* model is estimated using bivariate probit, the insignificant correlation between the *CEOATAX* and *BUATAX* models' error terms leads to a loss of efficiency consistent with the estimations reported in Greene (2003, 717). Following Greene (1998; 2003, 717), inferences concerning the effect of the *BUATAX* determinants on the probability that a firm compensates its BU managers on an after-tax basis should be based on the single

(*continued on next page*)

TABLE 6
BUATAX Model Estimations

$$BUATAX_i = \alpha_0 + \alpha_1 CEOATAX_i + \alpha_2 MULTI_i + \alpha_3 CAPINT_i + \alpha_4 LEV_i + \alpha_5 SIZE_i + \alpha_6 ROA_i + \alpha_7 VARROA_i + \alpha_8 GROWTH_i + \alpha_9 FWONLY_i + \alpha_{10} BUSTK_i + \gamma_{11-16} IND_i + \varphi_i$$

			(1)[a]		(2)[b]	
Independent Variable	**Coeff.**	**Pred. Sign**	**Coeff. Est.**	**p-value[c]**	**Coeff. Est.**	**p-value[c]**
Intercept	α_0	?	−1.418	.082	−1.421	.278
CEOATAX	α_1	+	0.945	.000	1.452	.230
MULTI	α_2	+	0.821	.101	0.707	.212
CAPINT	α_3	+	1.204	.048	0.928	.270
LEV	α_4	+	−1.371	.895	−1.285	.828
SIZE	α_5	+	0.099	.109	0.072	.324
ROA	α_6	+	−3.975	.959	−4.374	.942
VARROA	α_7	+	33.426	.305	24.341	.401
GROWTH	α_8	−	−0.592	.157	−0.526	.277
FWONLY	α_9	+	1.527	.000	1.484	.009
BUSTK	α_{10}	−	0.252	.854	0.242	.188
Model Chi-Squared (16 d.f.)			69.46		NA	
Pseudo R^2			.267		NA	

[a] Column (1) presents the results from a probit estimation of the *BUATAX* model.

[b] Column (2) presents results from the maximum likelihood bivariate probit estimation of the *CEOATAX* and *BUATAX* models. The correlation between the error terms of the *CEOATAX* and *BUATAX* selection equations is −.331 (p = .813).

[c] p-values are based on a one- (two-) tailed test where the coefficient sign is (not) predicted.

Variable Definitions:

BUATAX = 1 if the firm's BU managers are compensated on an after-tax basis;
CEOATAX = 1 if the CEO's annual bonus is based on after-tax accounting results;
MULTI = the ratio of foreign assets to the book value of total assets (*BVA*) as of year-end 1997;
CAPINT = the ratio of property, plant, and equipment to *BVA* as of year-end 1997;
LEV = the ratio of long-term debt to *BVA* as of year-end 1997;
SIZE = the natural logarithm of *BVA* as of year-end 1997;
ROA = the ratio of the average pre-tax income for the three years ending in 1997 to *BVA* as of year-end 1997;
VARROA = the variance in *ROA* for the three-year period ended in 1997;
GROWTH = the geometric mean annual growth rate of the market value of assets (*MVA*) for the three years 1995–1997;
FWONLY = 1 if the firm uses a firm-wide accounting measure, but not a unit-based measure, in its BU-manager incentive compensation plans;
BUSTK = 1 if stock price is used as a performance measure in the incentive compensation plans of a majority of the firm's business-unit managers; and
IND = vector of industry indicator variables for the mining and construction, transportation and communication, utilities, sales, financial institutions and insurance, and services industries (results not tabulated).

Footnote 24, continued
equation estimation of the *BUATAX* model, which is achieved by restricting the correlation in error terms between the *BUATAX* and *CEOATAX* equation to zero. As a sensitivity test, the *ETR* model was reestimated with the selectivity correction variables recomputed based on a bivariate estimation of the *CEOATAX* and *BUATAX* models in which the correlation parameter was restricted to zero. The results of this sensitivity test are substantially the same as those reported in Table 4.

Sensitivity Tests Focusing on High-MTR Firms

Evidence consistent with the hypothesis that accounting-based after-tax BU-manager performance measures lead to lower ETRs does not necessarily imply that such measures promote tax-planning effectiveness. Although low-MTR firms are expected to engage in tax-disadvantaged activities that would generally increase the ETR (Scholes et al. 2002), if after-tax performance measures lead these firms to engage in naïve tax-minimization strategies that lower the ETR, then these measures are not effective. To help rule out this behavior as a potential explanation and provide additional evidence concerning the effectiveness of after-tax BU-manager compensation, the tests are performed using only high-MTR firms, identified using the average Graham (1996) MTR for the three years ending in 1997. Using a subsample of firms with average MTRs in the top 75 percent of the MTR distribution (MTRs of 33.5 percent and higher), the *ETR* model estimation results are substantially the same as those reported in Table 4.[25] These results provide evidence that the negative relation between after-tax BU-manager compensation and ETRs is not a result of low-MTR firms' tax minimization strategies.

Sensitivity Tests Involving Firm Spending on the Tax Function and Outsourcing

To investigate differential effects of internal and external spending on the ETR, the *LSPEND* variable in the *ETR* model is first decomposed into two variables, *LSPENDOUT* and *LSPENDIN*, which represent outside and inside spending, respectively.[26] When the *ETR* model is re-estimated with these variables replacing *LSPEND*, the coefficient on *LSPENDIN* (−.782) is weakly significant in a one-tailed test ($p = .096$), while the coefficient on *LSPENDOUT* is insignificant, suggesting that only internal tax function spending is associated with lower ETRs. The coefficients on *BUATAX* and *CEOATAX* are substantially identical to those reported in Table 4.

Dunbar and Phillips (2001) find a negative relation between the use of after-tax operating-manager compensation and the proportion of corporate tax function expenditures made to external service providers. To investigate whether the proportion of outsourcing is a correlated omitted variable that could bias the *BUATAX* coefficient estimates reported in Table 4, *OUTSOURCE*, defined as the proportion of 1997 outside-to-total spending on the tax function, is added to the *ETR* model. When this revised *ETR* model is re-estimated, the coefficient on *OUTSOURCE* is .032 but insignificant ($p = .127$). The coefficient on *BUATAX* remains significantly negative (−.072, $p = .004$).

VI. CONCLUSION

This study investigates the relation between firms' effective tax rates (ETRs) and after-tax CEO and BU-manager accounting-based performance measures. The evidence is consistent with after-tax BU-manager performance measures leading to lower ETRs and economically significant tax benefits. This result highlights the importance of BU-manager involvement in tax-planning efforts that lower firms' ETRs and suggests that explicit accounting-based incentives are incrementally effective in promoting such efforts. The evidence does not support the hypothesis that after-tax CEO performance measures lead to lower ETRs. These results suggest that other incentives such as job retention are sufficient to motivate CEOs to focus on after-tax results, whereas annual after-tax accounting-based bonuses provide incremental motivation only for BU managers to consider their decisions' tax consequences. After-tax CEO performance measures could have an indirect negative effect on ETRs, however, because CEOs compensated after-tax are more likely to compensate their BU managers on an after-tax basis.

---

[25] In the estimation that includes firms with average MTRs greater than .335 ($n = 157$), the coefficient on *BUATAX* is −.080 and significant ($p = .005$).

[26] Mean (median) internal spending on the tax function, as a percentage of total SG&A, is .008 (.002). Mean (median) outsourcing expenditures as a percentage of total SG&A is .003 (.001). In t-tests of mean differences, neither internal nor external spending differs across types of bonus plans.

The results contribute to our overall understanding of the role that accounting-based incentive compensation plays in motivating managers' efforts. In particular, this finding adds to prior evidence (Guidry et al. 1999) that BU managers respond to unit-based incentives. This study also identifies the pre- versus after-tax nature of BU-manager compensation as a factor in firms' propensities to engage in tax-planning efforts. This evidence should assist future researchers in their investigations of effective tax planning strategies. Finally, the estimate of potential explicit tax benefits associated with compensating BU managers on an after-tax basis should be useful to corporate decision makers in their design of incentive compensation plans.

Interpretation of the results is subject to the following limitations. First, it is possible that for those firms not currently using an after-tax BU-manager performance measure, the associated costs of doing so may be greater than its expected benefits. Accordingly, the results do not lead to an inference that all firms should adopt this practice. Second, the financial-statement-based ETR does not reflect implicit taxes, defined as the decrease in pre-tax returns associated with tax-advantaged investments (Scholes et al. 2002). Tax-planning strategies that lower the ETR could also result in implicit taxes, reducing, at least in part, the explicit tax benefits obtained from using after-tax BU-manager compensation.

APPENDIX
SELECTED SURVEY INSTRUCTIONS AND QUESTIONS

General Instructions

This questionnaire includes questions about the compensation plans of your company's chief executive officer, tax department personnel, and business-unit managers. These questions relate to how these persons are compensated. There are no questions that request confidential compensation amounts. There are also questions concerning the use of certain corporate tax management practices. If you do not believe that you can answer questions that pertain to your company's tax department, please ask your company's top tax professional to complete this questionnaire.

For purposes of this study, two key terms are defined as follows:

1. The definition of a business unit follows the definition of an operating segment set forth in FAS No. 131, *Disclosures about Segments of an Enterprise and Related Information.* Accordingly a business unit is defined as a "component of an enterprise about which separate financial information [assets, liabilities, income, and expenses] is available that is evaluated regularly by the chief operating decision maker in deciding how to allocate resources and in assessing performance."
2. Your company is broadly defined and refers to the combined entity that is required to prepare and issue a consolidated financial statement; i.e., your company encompasses a parent company and all of the entities in which it has a 50 percent or more ownership interest. Some questions refer to specific individuals within your company (e.g., the CEO, CFO, the company's top tax professional, etc.). You should refer to individuals of the parent company in answering these questions. For example, "your company's CEO" refers to the CEO of the parent company, not the CEO of any of the company's particular business units or divisions.

All the information you provide is confidential and will be published only in summary, statistical form. You or your company will not be identified in any way. Accordingly, there are no foreseeable risks to you or your company. In order to get accurate information about corporate tax management practices and to properly assess the relationship between the use of these practices and companies' tax liabilities, we need information from all financial executives selected to participate in this study.

. . . .

Q-4 How many business units are there in your company's worldwide organization? *Estimates* of this number are acceptable.

____

Q-5 For the fiscal year ending in 1997, how much did your company spend, on a *worldwide basis*, for salaries and other compensation-related costs within the company and fees to outside consultants for tax compliance and tax-planning/consulting work? Please indicate amounts, in 000s, and the estimated percentages allocable to tax planning/consulting in the spaces provided. *Estimates* of these amounts and percentages are acceptable.

Outside Consults

	Total Expenditures	Estimated % Related to Tax Planning/Consulting
Within Company		
Outside Consults		

. . . .

Q-12 The CEO's annual bonus is based, at least in part, on an accounting measure (e.g., net income or loss, operating cash flow, etc.) computed (Circle one):

1 ON A PRE-INCOME TAX BASIS
2 ON AN AFTER-INCOME TAX BASIS
3 THE CEO'S ANNUAL BONUS IS NOT BASED, EVEN IN PART, ON AN ACCOUNTING MEASURE
4 NA—THE CEO DOES NOT RECEIVE AN ANNUAL BONUS

III. Business-unit managers generally receive incentive compensation based, at least in part, on four alternative "types" of performance measures: (1) the company's stock price (e.g., through the issuance of stock-based compensation such as stock options or stock appreciation rights, through cash/stock bonuses tied to the achievement of stock price levels and/or increases, etc.), (2) company-wide accounting measures (e.g., measures based on net income/loss, cash flow, etc.), (3) accounting measures computed at the business-unit level and/or group level (i.e., at a level other than on a company-wide basis), and (4) nonfinancial goals or measures (e.g., a performance appraisal rating). The following questions relate to the current use of the first three of these performance measures in the incentive compensation plans of a majority of your company's business-unit managers. In your responses to these questions, please consider the following:

1. If your company does not have two or more business units, please refer to the incentive compensation plans of a majority of your company's operating managers.
2. If a question asks whether a particular performance measure is being used in the incentive compensation plans of your company's business-unit managers, then please answer YES even if this measure is only one of two or more performance measures used in such plans.
3. The accounting-based measures addressed in these questions should be interpreted in a broad manner. Specific measures, such as Return on Assets (ROA) and Return on Equity (ROE), may be derived from these broadly defined accounting-based measures. For example, if your company uses ROA and computes this measure using pretax operating cash flow in the numerator, then this measure would be considered an accounting measure based on cash flow for purposes of this study.

Q-13 Is your company's *stock price* used as a performance measure in the incentive compensation plans of a majority of your company's business-unit (operating) managers (e.g., through the issuance of stock-based compensation such as stock options and stock appreciation rights, through cash/stock bonuses tied to the achievement of stock price levels and/or increases, etc.)? (Circle one)

1 YES
2 NO

Q-14 Is a *company-wide accounting measure* (e.g., a measure based on net income/loss, operating cash flow, etc.) used as a performance measure in the incentive compensation plans of a majority of your company's business-unit (operating) managers? (Circle one)

1 YES
2 NO

Q-15 The *company-wide accounting measure(s)* used in the incentive compensation plans of a majority of your company's business-unit managers is determined: (Circle one)

1 ON A PRE-INCOME TAX BASIS
2 ON AN AFTER-INCOME TAX BASIS

....

Q-17 Is a *business-unit level and/or group level (i.e., at a level other than on a company-wide basis) accounting measure* used as a performance measure in the incentive compensation plans of a majority of your company's business-unit managers? (Circle one)

1 YES
2 NO
3 NA—THE COMPANY DOES NOT HAVE TWO OR MORE BUSINESS UNITS.

Q-18 Which of the following statements best describes the extent to which the company's income tax expense affects the determination of the *business-unit and/or group level* *accounting measure(s)* used in the incentive compensation plans of *a majority* of your company's business-unit managers? (Circle one)

STATEMENT 1 THIS MEASURE IS DETERMINED STRICTLY ON A PRE-INCOME TAX BASIS. INCOME TAX EXPENSE IS *NOT* ALLOCATED TO THE COMPANY'S BUSINESS UNITS.

STATEMENT 2 THE COMPANY'S INCOME TAX EXPENSE IS ALLOCATED TO THE BUSINESS UNITS BASED SOLELY ON A CONSTANT COMPANY-WIDE PERCENTAGE (e.g., THE EFFECTIVE TAX RATE). THE EFFECTS OF INCOME-TAX-RELATED ITEMS ATTRIBUTABLE TO SPECIFIC BUSINESS UNITS ARE *NOT* CONSIDERED IN DETERMINING THIS MEASURE.

STATEMENT 3 THE EFFECTS OF A FEW (POSSIBLY ONLY ONE) INCOME-TAX-RELATED ITEMS ATTRIBUTABLE TO SPECIFIC BUSINESS UNITS ARE ALLOCATED TO THE BUSINESS UNITS IN DETERMINING THIS MEASURE.

STATEMENT 4 THE COMPANY'S TOTAL INCOME TAX LIABILITY IS EXTENSIVELY ALLOCATED AMONG THE BUSINESS UNITS IN DETERMINING THIS MEASURE.

• • • • • • • • • •

REFERENCES

Abarbanell, J., and B. Bushee. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research* 35 (Spring): 1–24.

Antle, R., and J. Demski. 1988. The controllability principle in responsibility accounting. *The Accounting Review* 63: 700–718.

Atwood, T. J., T. Omer, and M. Shelley. 1998. Before versus after-tax earnings as performance measures in compensation contracts. *Managerial Finance* 11: 30–44.

Bankman, J. 1994. The structure of Silicon Valley start-ups. *UCLA Law Review* 41 (September): 1737–1768.

Bushman, R., R. Indjejikian, and A. Smith. 1995. Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies. *Journal of Accounting Research* 33 (Supplement): 101–128.

Carnes, G., and D. Guffey. 2000. The influence of international status and operating segments on firms' choice of bonus plans. *Journal of International, Accounting, Auditing, and Taxation* 9 (1): 43–57.

Dhaliwal, D., J. Sneed, and R. Trezevant. 2000. Factors associated with the choice to use before-tax or after-tax earnings as the contracting variable in annual bonus plans. Working paper, University of Southern California.

Dillman, D. 1978. *Mail and Telephone Surveys: The Total Design Method.* New York, NY: John Wiley.

Douglas, R., P. Ellingsworth, and J. McAndrews. 1996. What management expects: Evaluation of the tax department and its executives. *The Tax Executive* (Jan/Feb): 29–32.

Dunbar, A., and J. Phillips. 2001. The outsourcing of corporate tax function activities. *The Journal of the American Taxation Association* 23 (Fall): 35–49.

Ernst & Young. 1999. Transfer pricing 1999 global survey: practices, perceptions, and trends. *Tax Notes* (November 22): 1073–1100.

Graham, J. 1996. Proxies for the corporate marginal tax rate. *Journal of Financial Economics* 42: 187–221.

Greene, W. 1998a. Gender economics courses in liberal arts colleges: Further results. *Journal of Economic Education* 29 (Fall): 291–300.

———. 1998b. *LIMDEP Version 7.0 User's Manual.* Plainview, NY: Econometric Software, Inc.

———. 2003. *Econometric Analysis.* 5th edition. Upper Saddle River, NJ: Prentice Hall, Inc.

Guenther, D., and D. Jones. 2002. Valuation implications of changes in firms' effective tax rates. Working paper, University of Colorado at Boulder.

Guidry, F., A. Leone, and S. Rock. 1999. Earnings-based bonus plans and earnings management by business-unit managers. *Journal of Accounting and Economics* 26: 113–142.

Gupta, S., and K. Newberry. 1997. Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data. *Journal of Accounting and Public Policy* 16: 1–34.

Hanlon, M. 2002. The persistence and pricing of earnings, accruals, and cash flows when firms have large book-tax differences. Ph.D. dissertation, University of Washington.

———, and T. Shevlin. 2002. Accounting for tax benefits of employee stock options and implications for research. *Accounting Horizons* 16 (1): 1–16.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7: 85–107.

Heckman, J. 1976. The common structure of statistical models with truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5: 475–492.

———. 1990. Varieties of selection bias. *American Economic Review* (Papers and Proceedings) 80: 313–318.

———. 1997. Instrumental variables. A study of implicit behavioral assumptions used in making program evaluations. *Journal of Human Resources* 32: 441–462.

Ittner, C., and D. Larcker. 2001. Assessing empirical research in managerial accounting: A value-based management perspective. *Journal of Accounting and Economics* 32 (December): 349–410.

Kahn, L., and P. Sherer. 1990. Contingent pay and managerial performance. *Industrial and Labor Relations Review* 43 (February): 107S–120S.

Keating, S. 1997. Determinants of divisional performance evaluation practices. *Journal of Accounting and Economics* 24: 243–273.

Krull, L. 2001. The relation between permanently reinvested earnings, repatriation taxes, and earnings management incentives of U.S. multinationals. Ph.D. dissertation, The University of Arizona.

Larcker, D. 1983. The association between performance plan adoption and corporate capital investment. *Journal of Accounting and Economics* 5: 3–30.

Lev, B., and R. Thiagarajan. 1993. Fundamental information analysis. *Journal of Accounting Research* 31 (2): 190–215.

McLemore, I. 1997. World-class tax planning. *Controller Magazine* (April): 41–44.

Milgrom, P., and J. Roberts. 1992. *Economics, Organization and Management.* Englewood Cliffs, NJ: Prentice Hall.

Mills, L., and M. Erickson, and E. Maydew. 1998. Investments in tax planning. *The Journal of the American Taxation Association* 20 (Spring): 1–20.

Mintz, S. 1999. A taxing challenge. *CFO Magazine* (November).

Nelson, F. 1984. Efficiency of the two-step estimator for models with endogenous sample selection. *Journal of Econometrics* 24: 181–196.

Newman, H. 1989. Selection of short-term accounting-based bonus plans. *The Accounting Review* 64: 758–772.

Scholes, M., M. Wolfson, M. Erickson, E. Maydew, and T. Shevlin. 2002. *Taxes and Business Strategy: A Planning Approach.* 3rd edition. Upper Saddle River, NJ: Prentice Hall, Inc.

Swenson, C. 1999. Increasing stock market value by reducing effective tax rates. *Tax Notes* (June 7): 1503–1505.

Tunali, I. 1986. A general structure for models of double-selection and an application to a joint migration/earnings process with remigration. *Research in Labor Economics* 8B: 235–282.

Vella, F. and M. Verbeek. 1999. Estimating and interpreting models with endogenous treatment effects. *Journal of Business and Economic Statistics* 17 (4): 473–478.

Wales, T. and A. Woodland. 1980. Sample selectivity and the estimation of labor supply functions. *International Economic Review* 21: 437–468.

Wallace, J. 1997. Adopting residual income-based compensation plans: Do you get what you pay for? *Journal of Accounting and Economics* 24: 275–300.

Wang, S. 1991. The relation between firm size and effective tax rates: A test of firms' political success. *The Accounting Review* 66 (January): 158–169.

Wilkie, P. 1988. Corporate average effective tax rates and inferences about relative tax preferences. *The Journal of the American Taxation Association* 10 (Fall): 75–88.

Wilson, G. P. 1995. Corporate tax effectiveness: Adding value through the tax function. Working paper, Harvard University.

Ziegler, B. 1997. IBM's profit tops analyst expectations. *Wall Street Journal* (July 22): A3.

THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 875–902

The Contribution of Fundamental Analysis after a Currency Devaluation

Edward P. Swanson
Lynn Rees
Texas A&M University

Luis Felipe Juarez-Valdes
Universidad de las Americas-Puebla

ABSTRACT: For a sample of companies traded on the Mexican Bolsa, fundamental analysis is used to investigate the value of financial statement information to investors after the December 1994 currency devaluation. Associations with contemporary returns show that earnings in the year of the devaluation lose value relevance, but fundamental signals, which incorporate the more detailed accounting information provided in financial statements, retain considerable explanatory power (R^2 is 25 percent). After the devaluation, fundamental signals based on changes in selling and administrative expenses and changes in gross margin are significant in several analyses, including predictions of future earnings, analysts' forecast revisions, and analysts' forecast errors. Because analysts underutilize those signals, an opportunity exists after the devaluation for a substantial profit from a zero investment trading strategy.

Keywords: *currency devaluation; financial statement analysis; Mexico; economic shock.*

Data Availability: *All data are available from public databases identified in the paper.*

I. INTRODUCTION

Investors' need for forward-looking accounting information is greatly increased when an economic shock occurs during a reporting period. This is particularly true when the shock occurs late in the reporting period, so that current earnings information cannot be extrapolated to the future. Under these conditions, earnings lose value relevance and investors are forced to rely on other information. The basic idea underlying this paper is

---

We acknowledge Brian Bushee, Bill Cready, Mahendra Gujarathi, Don Herrmann, Mary Lea McAnally, Gary Meek, Grace Pownall, Mike Wilkins, two anonymous reviewers, and participants at workshops at Oklahoma State University, Texas A&M University, and University of North Texas for helpful comments on earlier drafts. We also thank C. P. Francisco A. Castor del Valle of KPMG Cardenas Dosal, S. C. Peat Marwick, and C. P. Javier Cocina Martinex of Instituto Mexicano de Contadores Publicos, A.C., for meeting with the authors to answer questions about the Mexican accounting system. The Center for International Business Studies at Texas A&M University provided summer funding. We gratefully acknowledge the contribution of I/B/E/S International Inc. for providing earnings per share forecast data, available through the Institutional Brokers Estimate System. This data has been provided as part of a broad academic program to encourage earnings expectations research.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted January 2002
Accepted January 2003

that, when the value-relevance of earnings is reduced by severe economic change, the detailed performance information provided in financial statements may still be useful in predicting future earnings and cash flows. Currency devaluations provide a natural experimental setting to investigate the effects of severe economic change. The findings from this setting likely extend to other instances in which firms encounter an economic shock during a reporting period (e.g., a terrorist attack, an oil embargo, or a labor strike).

The specific setting we investigate is the December 1994 peso devaluation in Mexico. The economic shock from the peso devaluation was typical of other currency devaluations. Inflation increased dramatically from 7 percent in calendar 1994 to 52 percent in 1995, and stock prices dropped by about 50 percent (see Figure 1, Panel A). A unique aspect of the 1994 Mexican devaluation is that most of the exchange rate decline occurred during the final week of December. The timing of the peso devaluation produces an ideal research setting. First, 1994 earnings clearly cannot be extrapolated to the future, so investors need alternative information. Second, we can use annual report data to test whether pre-shock accounting information can provide forward-looking information about post-shock operating performance and stock returns.

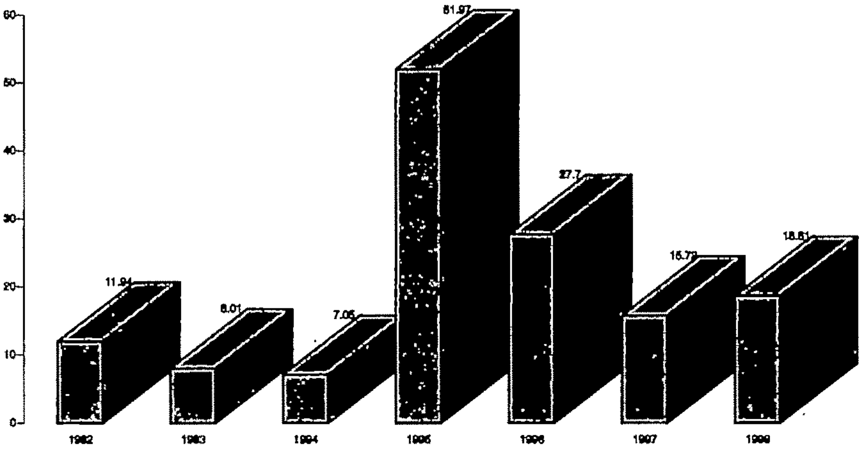
The primary finding to-date from research on currency devaluations is that earnings lose their value relevance (Graham et al. 2000; Ho et al. 2001; Davis-Friday and Gordon 2002). This study differs from prior research by using the fundamental analysis approach developed by Lev and Thiagarajan (1993) (hereafter LT) as an alternative way to investigate the value relevance of accounting information after a devaluation. Fundamental analysis uses the detailed information provided in financial statement line items to provide a finer analysis of a firm's performance than can be obtained using a summary measure, such as earnings. By considering changes in key components of operating performance (e.g., changes in gross margin or changes in expenses relative to sales), fundamental analysis has the potential to capture more completely than earnings the value relevance of accounting information after a macroeconomic shock.

The sample consists of companies that traded on the Mexican Bolsa during the period from 1992 to 1998. The results begin with 1993, however, since we use 1992 data to calculate changes. We find that annual earnings levels have considerable value relevance during the non-devaluation years and including the fundamental signals adds additional explanatory power. In contrast, earnings are not value relevant in the 1994 devaluation year. Importantly, a model that includes both the fundamental signals and earnings has considerable value relevance in 1994 (R^2 of 25 percent), primarily due to the selling and administrative (S&A) expense signal. Additional tests using interim accounting disclosures show that earnings remain value relevant during the first three quarters of 1994. In the fourth (devaluation) quarter, earnings lose value relevance, but fundamental analysis, primarily the selling and administrative expense signal, results in significant value relevance (R^2 of 24.7 percent). The tests using interim information indicate the increased contribution of fundamental analysis is due to the devaluation event, rather than a gradual decline in economic conditions throughout 1994.

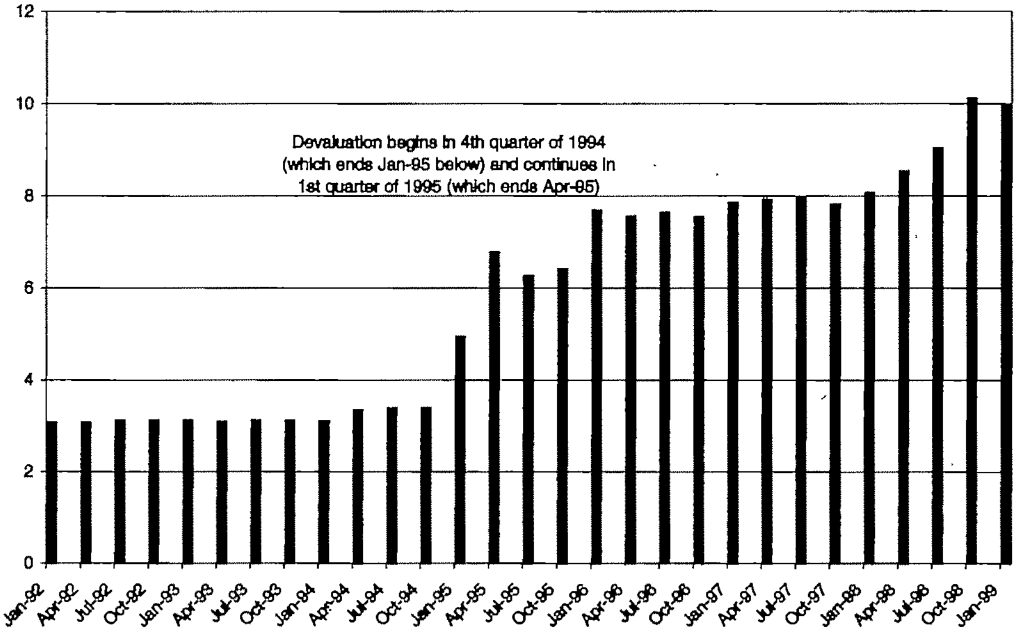
To provide further information about whether the detailed accounting data used in fundamental analysis provide forward-looking information in 1994, we conduct several additional analyses. These analyses do not rely on an assumption that contemporaneous market reactions capture the full extent to which accounting information may be of value to investors. Using the approach in Abarbanell and Bushee (1997), we find that fundamental signals based on 1994 financial statements predict one-year-ahead earnings changes, analysts' forecast revisions, and analysts' forecast errors. This finding indicates that the signals are useful in predicting future earnings and that analysts use the signals in revising forecasts,

FIGURE 1
Inflation and Exchange Rate Statistics (per Banco de Mexico)

Panel A: Annual Inflation Rates (Mexican CPI)



Panel B: Number of Pesos Required to Obtain One U.S. Dollar from Jan. 92 to Jan. 99



but their revision does not fully reflect all the information about future earnings. In particular, two fundamental signals are consistently significant in the analyses: selling and administrative expense, and gross margin. Companies that are able to control the level of selling and administrative expenses relative to sales provide higher earnings after the devaluation, so Mexican companies must pay particular attention to controlling overhead costs. The rationale is that such costs are sticky, tending to increase more with an economic upturn than the decrease with a downturn (Anderson et al. 2002). In addition, firms that are able to maintain or increase their gross margin provide significantly higher future earnings. These firms are better able to pass on to their customers the higher costs from the increase in inflation. Tests using quarterly accounting information and forecast revisions immediately before and after the devaluation indicate that the increased contribution of the fundamental signals is due to the devaluation event, rather than to a gradual decline in economic conditions throughout 1994.

Because the analysis of forecasts provides evidence that analysts underutilize the information contained in fundamental signals, we then examine if a profitable trading strategy can be developed. Using the approach in Abarbanell and Bushee (1998), we construct a hedge portfolio in 1994 and document abnormal returns from a trading strategy based on the selling and administrative expense, and gross margin fundamental signals. This trading strategy does not produce significant abnormal returns in years other than 1994. We conclude from these various tests that fundamental analysis assumes additional importance in the wake of a macroeconomic shock but the market does not fully impound this fact.

The findings from this study are of potential interest to several groups. Accounting academics are interested in the general topic of how accounting information is priced by the market. The finding that the incremental value relevance of detailed financial statement information (about changes in assets, liabilities, revenues, and expenses) over that provided by earnings is increased by an economic shock is a new insight into how accounting information is priced. Investors should value knowing the economic conditions under which fundamental analysis provides information that is incremental to earnings. Investors, especially hedge fund managers, should be interested in the finding that the market does not fully impound the information in specific signals, since they may be able to develop a profitable trading strategy by going long and short on stocks, based on fundamental signals. Finally, accounting standard-setters, such as the FASB, International Accounting Standards Board, and the U.K.'s Accounting Standards Board, may be interested in our results for their project on reporting performance, which is considering the predictive value of various financial statement displays (see http://www.FASB.org for a project summary and related links).

The remainder of the paper consists of six sections. The second section describes the source of the data and the accounting system in Mexico. The third section discusses the research methodology. The fourth section provides a summary of the fundamental signals and their hypothesized effects in Mexico. The fifth section presents our empirical results. The final section provides a summary and conclusion.

II. DESCRIPTION OF DATA SOURCE AND MEXICAN ACCOUNTING SYSTEM

The financial statement and stock price data used in this study were obtained from the Economatica database, which provides accounting data and stock returns for Latin American companies. Economatica reports accounting information based on the standardized format used in filing with the Mexican Bolsa. This format provides the intermediate components of financial statements that are needed for fundamental analysis. The sample consists of all companies that traded on the Mexican Stock Exchange during the period 1992–1998 (the 1992 year is used to calculate changes). The sample period begins with 1992 because this is the first year for which the Economatica database provides an extensive sample.[1]

In accord with Bulletin B-10 (Instituto Mexicano de Contadores Publicos [IMCP] 1984), *Recognition of the Effects of Inflation in Financial Information*, Mexican companies have reported replacement-cost, price-level-adjusted (RCPLA) information in the primary financial statements since 1984 (under an evolving set of rules). Under RCPLA, monetary assets and liabilities for the latest year do not require restatement because they are already stated in terms of year-end purchasing power. However, companies do report purchasing power gains and losses on monetary assets and liabilities held during the year as a component of earnings. Inventory and property, plant, and equipment are revalued at year-end and the holding gain/loss is reported in stockholders' equity. Through 1996, inventory was reported at estimated cost to manufacture or repurchase at year-end prices, and appraisals were used to value property, plant, and equipment at replacement cost. Beginning January 1, 1997, Mexican companies have been required under the Fifth Amendment to Bulletin B-10 to use the consumer price index to revalue all nonmonetary items, including inventory and property, plant, and equipment. Valuations during our sample period, therefore, reflect a mix of replacement cost and general-price-level-adjusted amounts. (For expository convenience, we use the acronym RCPLA for all periods.)

Under RCPLA accounting, amounts for every time period are reported on the financial statements in purchasing power at the year-end of the most recent reporting period. Assets and liabilities for the current year are already valued in terms of year-end purchasing power (as discussed above). Operating results during the current year and financial statements from prior years need to be restated. The Mexican Consumer Price Index is used as the measure of purchasing power. In accord with the Third Amendment of Bulletin B-10, which became effective in 1990, the restatement procedure works as follows: If inflation during the current year has been 30 percent, then every amount from the prior year's financial statement is multiplied by 1.3 to obtain the balances used for comparative purposes in the current year's statement. Revenues and expenses that occur during the current year are rolled forward into year-end prices using monthly price indexes. If inflation occurs evenly throughout the year, then the monthly adjustments will often approximate the results that would be obtained by multiplying revenues and expenses by an average index for the year. For 30 percent inflation, the average index would be 1.15. The primary purpose of the restatement is to facilitate the analysis of year-to-year changes in operating results.[2] Since fundamental signals are calculated using changes in income statement amounts, assets, and liabilities as reported in the annual report (e.g., change in selling and administrative expenses, change in accounts receivable), the signals used in this study are calculated in constant, consumer purchasing power. This paper is the first to study fundamental signals based on RCPLA accounting data.

---

[1] In using the Economatica database, researchers need to make two types of adjustments to the data. First, the signs of certain earnings components are inconsistent. For example, a positive number for net interest expense sometimes represents an expense that is subtracted from net income and other times represents net interest income that is added to net income. We adjusted for this inconsistency by downloading from Economatica various income subtotals, which allowed us to determine the true sign of an income statement component. Second, the inflation adjustment factor applied to financial statement items is sometimes inconsistent. In particular, we discovered that for most (but not all) of the companies in our sample, the 1994 data were adjusted using the 1995 inflation-adjustment factor; this problem exists to a much lesser extent in other years (we are grateful to an anonymous reviewer for pointing this problem out to us). We made corrections by downloading from Economatica both the reported and inflation-adjusted financial statement numbers and comparing them to determine the inflation-adjusted factor that was actually used by Economatica. We adjusted the data using a consistent factor. In a few instances, we determined the nature of the amounts reported by referring to the Worldscope database.

[2] Readers interested in an example of RCPLA accounting should see Gordon (2001, Appendix). Figure 1, Panel A, summarizes the magnitude of price changes over our test period.

III. RESEARCH METHODOLOGY

Association of Earnings and Fundamental Signals with Contemporaneous Returns

The value relevance of earnings is assessed by estimating the following equation:

$$R_{it} = \alpha_0 + \beta_1 E_{it} + \nu_{it} \tag{1}$$

where R represents a market-adjusted, 12-month return, ending three months following the fiscal year-end (which is December for all firms on the Bolsa); E is earnings-per-share deflated by price at the beginning of the return accumulation period; i and t are firm and time subscripts, respectively; and υ is an error term assumed iid $N(0,\sigma^2)$.[3] Abnormal returns were calculated as the difference between the firm's raw return and a market-adjusted amount using the IPC (*Prices and Quotations Index*) from the Bolsa. Model (1) is estimated for the 1994 devaluation year and again for the five non-devaluation years.

As discussed further in the next section of the paper, we use the fundamental signals originally developed by LT (with the addition of a leverage signal). The fundamental signals model is expressed as follows:

$$R_{it} = \alpha_0 + \beta_1 PTE_{it} + \sum_{j=2}^{k} \beta_j X_{jit} + \omega_{it} \tag{2}$$

Similar to LT, PTE is defined as pre-tax earnings multiplied by 1 minus the tax rate in year $t-1$ and deflated by price at the beginning of the return accumulation period. X_j represents the jth fundamental signal that is included in the model. Model (2) is also estimated separately for the 1994 devaluation year and over the five non-devaluation years.

We investigate the value relevance of earnings and fundamental signals in three steps. First, we use Model (1) to determine whether earnings are value relevant in the 1994 devaluation-reporting period and in the non-devaluation-reporting periods. Second, we use Model (2) to determine whether the fundamental signals provide value relevance that is incremental to earnings in the devaluation and non-devaluation periods. Third, we examine whether the coefficients for earnings and the fundamental signals differ between the devaluation and non-devaluation periods.

Any increase in value relevance for the fundamental signals when comparing 1994 to the non-devaluation annual periods could be due to the December devaluation event or economic conditions during the year that triggered the devaluation. To investigate this issue, we use interim accounting data to fit Models (1) and (2) for two subperiods of 1994. The first subperiod uses deflated earnings for the first three calendar quarters of the year regressed on market-adjusted returns beginning April 1, 1994 (so that calendar 1993 earnings have been released) and extending through November 30, 1994 (so that third quarter 1994 earnings have been released).[4] The second subperiod uses fourth quarter earnings regressed on returns surrounding the devaluation event, covering December 1, 1994 through March 31, 1995 (see Figure 1, Panel B). We then investigate whether the results found in the comparison of Models (1) and (2) for 1994 is due to the fourth (devaluation) quarter, the first three quarters, or both.

---

[3] We also examined earnings/returns models that included as independent variables only deflated earnings change and deflated earnings change together with deflated earnings level. Consistent with Gordon (2001, 193), we find that earnings changes add little explanatory power.

[4] We obtained the 1994 third-quarter earnings announcement date for all companies where this variable is available on I/B/E/S. Although earnings announcement dates are available in this quarter for only 15 companies, every announcement was made prior to November 30, 1994.

Association of Fundamental Signals with Future Earnings Changes, Analysts' Forecast Revisions, and Analysts' Forecast Errors

The notion underlying fundamental analysis is that it provides insight concerning the quality of earnings and, therefore, provides information about the probability that current earnings will persist into the future. Although an association between fundamental signals and contemporaneous returns indicates that the fundamental signals are capturing information used by market participants, it does not necessarily mean that the signals themselves are useful in assessing earnings quality. To provide this information, we apply the methodology used by Abarbanell and Bushee (1997) (hereafter AB). First, we examine if an association exists between the fundamental signals and future earnings changes. A significant association indicates that the signals provide insight into earnings persistence. Second, we examine the relation between the fundamental signals and revisions in analysts' earnings forecasts to determine whether analysts use the information provided by the signals in revising their forecasts. Third, we examine the relation between the fundamental signals and analysts' earnings forecast errors to determine whether analysts completely use the signals in revising their forecasts.

We examine the usefulness of the fundamental signals in Mexico in predicting future earnings changes, analysts' forecast revisions, and analysts' forecast errors using the following equation:

$$Y_{it} = \alpha_0 + \beta_1 PTE_{it} + \sum_{j=2}^{k} \beta_j X_{jit} + \omega_{it} \tag{3}$$

where Y represents the change in one-year-ahead earnings (denoted ΔE_{it+1}), the revised analyst forecast for year $t+1$ after year t's earnings are announced (denoted REV_{it}), or the analyst forecast error for year $t+1$ (denoted AFE_{t+1}). ΔE_{it+1} is defined as earnings-per-share in year $t+1$ minus earnings-per-share in year t deflated by price at the beginning of the returns accumulation period. Similar to Abarbanell and Bushee (1997), REV_{it} is defined as follows:

$$REV_{it} = [(PostF_{t+1} - E_t) - (PreF_{t+1} - F_t)]/P_{t-1} \tag{4}$$

where $PostF_{t+1}$ is the first available mean consensus forecast for earnings in year $t+1$ issued subsequent to the year t earnings announcement, E_t is realized earnings in year t, $PreF_{t+1}$ is the first available mean consensus forecast in year t for earnings in year $t+1$, and F_t is the first available mean consensus forecast in year t for year t earnings.[5] REV_{it} can be thought of as the forecast revision ($PostF_{t+1} - PreF_{t+1}$) after removing the effect of the current period's forecast error on the revised forecast of future earnings (i.e., removing $E_t - F_t$). Finally, AFE_{it+1} is defined as realized earnings-per-share in year $t+1$ minus the first forecasted earnings-per-share for year $t+1$ issued subsequent to year t's earnings announcement, deflated by price at the beginning of the returns accumulation period.

---

[5] All data to construct ΔE_{it+1}, AFE_{it+1}, and REV_{it} were obtained from I/B/E/S to ensure that we use the realized earnings construct that analysts are trying to forecast. We do not adjust these dependent variables to a constant purchasing power. Deflating by stock price at the end of March reduces the need to do so. In addition, analysts' earnings forecasts would include anticipated inflation, so an additional adjustment would double-count inflation.

A significant coefficient on a fundamental signal indicates the signal is useful in predicting the dependent variable (either realized earnings changes, analysts' forecast revisions, or analysts' forecast errors). Comparing the association of a fundamental signal across the three equations can provide insight as to whether analysts fully use the information about future earnings contained in the signal. For example, if a fundamental signal is associated with future earnings changes and analysts' forecast errors but not with analysts' forecast revisions, then this finding would suggest that analysts do not use the information contained in the signals. Underutilization suggests that investors could take advantage of analysts' inefficiencies to earn abnormal profits.

IV. FUNDAMENTAL SIGNALS

We relied primarily on the fundamental signals chosen by LT, who based their choice on an examination of the types of accounting information used by financial analysts in the U.S. Before using the LT fundamental signals, however, we examined sell-side analyst reports for seven large Mexican firms. We found that the report format and types of analyses are very similar to that for U.S. companies. Sufficient information is available to calculate five fundamental signals that were statistically significant in the LT model. Although several signals cannot be used, the tests in this study include most of the fundamental signals that were found to be significant in the LT empirical analyses.[6] In a more recent study, Amir et al. (1999) also use only these five signals.

The five signals were designed by LT so that a negative sign reflects good news (and, therefore, positive returns) under the most common interpretation by analysts for U.S. firms. Under the economic conditions faced by firms in Mexico, the expected sign can differ from the U.S. Each of the five signals is discussed briefly in the following paragraphs, with comments on possible differences in applying the signals to Mexican companies. We also discuss a borrowing cost signal, which could be important to Mexican companies because they face volatile interest rates.[7]

1. Inventory: An increase in inventory relative to sales is generally interpreted by financial analysts as a negative signal for two reasons. First, such an occurrence indicates a greater chance that inventory will become obsolete. Second, holding costs are an increasing function of the amount of inventory on hand. On the other hand, a higher level of inventory might be viewed as a positive signal, since it reduces the chances of experiencing an inventory shortage and could signal that managers expect an increase in future sales. In an inflationary economy, an increase

---

[6] LT report their results using the average for the last two years as a measure for the expected level of each variable. LT (1993, 194) state that expectation models using just the prior year yield very similar results. This study uses the single year (random walk) formulation to avoid losing another year of results. This practice allows us to include new firms in the second year, rather than the third year, in which they are included in the Economatica database.

[7] In earlier analyses, several other variables that are unique to Mexico and/or might be expected to be important to investors during a currency devaluation were included in the full model. These variables included variations of foreign debt as a percentage of total debt, foreign sales as a percentage of total sales, holding gains/losses (which are reported by Mexican companies in stockholders' equity), and integral cost of financing (which is an income statement item that consists of the sum of purchasing power gains/losses, exchange rate gains/losses, and net interest). None of these variables provided a consistent improvement for our tests and separate reporting of the integral cost of financing resulted in multicollinearity problems in several tests. Thus, these variables are not reported in our final models.

in inventory may also be a positive sign if the additional inventory has been obtained or produced in anticipation of cost increases. Further, inventory purchases, rather than cost of goods sold, are deducted for taxes in Mexico. As a result, while a negative sign is expected in the U.S., an increase in inventory could be either a positive or negative signal in Mexico. The signal is computed as follows:

$$Inv = \frac{Inventory_t - Inventory_{t-1}}{Inventory_{t-1}} - \frac{Sales_t - Sales_{t-1}}{Sales_{t-1}}$$

2. Accounts Receivable: Disproportionate increases in accounts receivable relative to sales are considered an unfavorable signal since this could indicate collections are later in relation to the firm's credit policy. This practice is more likely to occur in inflationary periods because customers benefit more by delaying payments. Higher receivables could also indicate that the company has relaxed credit terms in order to sustain a given level of sales. Last, a significant increase in accounts receivable might indicate the company has managed earnings by "stuffing the channel" by shipping goods near the end of a reporting period. In each case, consistent with the U.S., a negative association is expected for Mexico between this fundamental signal and returns. The accounts receivable signal is computed as:

$$AR = \frac{Acct.Rec._t - Acct.Rec._{t-1}}{Acct.Rec._{t-1}} - \frac{Sales_t - Sales_{t-1}}{Sales_{t-1}}$$

3. Gross Margin: A disproportionate (to sales) decrease in the gross margin amount (sales minus cost of goods sold) is viewed as a negative signal by analysts about the long-term performance of the firm. A decrease in the gross margin signal could be caused by a decrease in sales price or an increase in input costs. A decrease in sales price could reflect increased competition or customer resistance to prices. An increase in input costs could be due to fixed or variable costs. Cost increases are more likely when inflation is high, a more frequent condition in Mexico than in the U.S. In each case, the permanence of future earnings is likely to be impaired by a decline in gross margin that is not offset by an increase in sales, so we expect a negative sign. The signal is defined as follows:

$$GM = \frac{Sales_t - Sales_{t-1}}{Sales_{t-1}} - \frac{Gross\ Margin_t - Gross\ Margin_{t-1}}{Gross\ Margin_{t-1}}$$

4. Selling and Administrative Expenses: An increase in administrative expenses relative to sales can indicate a loss of control over fixed expenses that cannot be passed on to customers and will adversely affect future cash flows. In addition, an increase in selling expenses relative to sales can indicate that an increased effort to produce sales was not entirely effective. Recent research has shown administrative expenses to be "sticky," meaning they tend to increase more with economic upturns than the decrease with a downturn (Anderson et al. 2002). While LT found this signal to be important in the U.S., we expect it could be even more important in Mexico because the level of sales can change substantially from period to period. The S&A fundamental signal, which is expected to have a negative sign, is defined as:

$$S\&A = \frac{S\&A_t - S\&A_{t-1}}{S\&A_{t-1}} - \frac{Sales_t - Sales_{t-1}}{Sales_{t-1}}$$

where *S&A* represents selling and administrative expenses.

5. Effective Tax Rate: A change in the effective tax rate that is due to reasons other than a statutory tax rate change is considered by analysts to be transitory in nature. The effective-tax-rate fundamental signal is calculated by splitting the change in earnings into two components. The first component is the current year's after-tax earnings that would have occurred if the prior year's tax rate had been in effect. The second component, the tax signal component, is the effect of the tax rate change on the current level of earnings. Using mathematical notation, the expression is:

$$E_t = PTE(1 - T_{t-1}) + PTE(T_{t-1} - T_t).$$

LT predict and find a negative relation between this fundamental signal and returns. The expected sign is more complicated for Mexican companies since a change in effective tax rate is more likely to have a permanent effect on earnings. Mexican businesses pay the greater of an income tax or an asset tax, resulting in taxes due for most loss years. If the tax rate increases because a firm paid an asset-based tax in the prior year, but now owes an income tax in the current year, then the signal could capture a permanent effect on earnings.[8] As a result, the tax signal could be positive or negative for Mexican companies.

6. Borrowing Costs: Leverage is a widely used ratio in fundamental analysis (Penman 2001), although LT did not include it among their set of fundamental signals. A firm that earns a higher (lower) rate of return on invested capital than the rate of interest paid on debt provides a higher (lower) return to stockholders. However, borrowing costs are highly volatile in Mexico, regardless of whether a firm borrows domestically or from outside the country. If a company has debt denominated in pesos, then domestic interest rates can increase dramatically to reflect higher anticipated inflation. If a Mexican firm has debt denominated in the U.S. dollar or another foreign currency, then the amount of pesos required for interest and principal payments increase immediately after a currency devaluation. Because borrowing costs are so volatile in Mexico, the fundamental signal is expected to have a negative sign. It is defined as:

$$Lev = \frac{Total\ Liabilities}{Total\ Assets}$$

---

[8] Interpretation of the tax signal is even more complicated because of the existence of a provision that allows recovery under some circumstances of the amount by which an asset tax exceeded the amount that would have been paid under an income tax that year. And during our sample period, the carry-forward recovery period was changed from five to ten years. The asset tax is roughly 2 percent of assets, although some liabilities can be deducted from the asset base (but not debts owed to foreign entities).

V. RESULTS

Descriptive Statistics and Benchmark Model

The Economatica database provides a total of 444 firm-year observations with the requisite financial statement and returns data, with the number of observations increasing over time. Given the volatile economic environment of Mexico during our sample period, it is not surprising that a preliminary examination of the data revealed some extreme fluctuations in returns, earnings, and other variables employed in our regression models. To control for influential observations, we winsorized all regression variables at the 5 and 95 percent tails of their respective distributions.

Table 1, Panel A, presents descriptive statistics for each year from 1993 to 1998, with amounts that are statistically significant at the .05 level presented in boldface. Since the focus of the paper is on the effects of the 1994 devaluation, the following discussion considers how 1994 differs from the other years. For the 1994 devaluation year, the means and medians for pretax earnings deflated by price (*PTE*) are not significantly different from zero.[9] In contrast, *PTE* is significantly positive in the other years with the exception of the mean for 1998, a year in which the peso exchange rate declined due to uncertainty generated by the Asian currency crisis (see Figure 1, Panel B). With regard to the fundamental signals, the inventory (*Inv*) and accounts receivable (*AR*) signals in 1994 are both significantly positive and the magnitudes are unusually large, indicating that inventory and accounts receivable grew faster than sales (i.e., downbeat signals). In 1995, the signs change to significantly negative as Mexican companies responded to the crises by reducing the level of inventory and accounts receivable. A similar reduction continued in 1996, which was then followed by increases in 1997. Leverage (*Lev*), which is always significantly different from zero, rose in the 1994 devaluation year and remained high in 1995 before declining in the following years. The higher leverage likely results from an increase in the number of pesos required to settle debt denominated in the U.S. dollar (or another foreign currency).

The magnitude of the correlations in the matrix in Table 1, Panel B, suggests that multicollinearity among the fundamental signals will not pose a significant problem. The top line of Table 1, Panel B, provides a univariate test of associations with returns for each of the explanatory variables using data pooled over the 1993 to 1998 period. As expected, pretax earnings (*PTE*) has a significant, positive association with returns (*R*). Three of the fundamental signals, gross margin (*GM*), selling and administrative expense (*S&A*), and leverage (*Lev*), are significantly negatively correlated with returns (as hypothesized).

Table 1, Panel C presents descriptive evidence about the properties of analysts' forecasts in Mexico. Analysts overestimated earnings (*FE*) by a median percentage of a little over 4 percent in 1994 and 1998. Analysts may not have anticipated the decline in the peso exchange rate in 1994 and 1998 (see Figure 1, Panel B) and the resultant negative effect on reported earnings. In general, negative signs indicate an overestimation of earnings, and research has found this occurs in several countries (e.g., Brown et al. 1985; O'Brien 1988; Francis and Philbrick 1993; Capstaff et al. 1998). Analysts' consensus forecast revisions (*REV*) measure the forecasted change in earnings from year t to $t+1$ after controlling for the current period's earnings surprise. The significant positive amounts for *REV* in 1994 are influenced by large negative forecast errors (*FE*) in that year (see Equation (4)).

The dispersion in analysts' forecasts is measured as the standard deviation of all available forecasts that comprise F_t divided by F_t. Dispersion increases significantly in 1995 and 1996 and then gradually returns to more normal levels, indicating that the 1994 devaluation reduced the predictability of earnings in subsequent years. The mean number of analysts following a company, which also corresponds to the number of forecasts that

---

[9] We use the expression "pretax earnings" to be consistent with LT's terminology, but this earnings measure has been reduced by taxes calculated using the tax rate from the prior year.

TABLE 1
Descriptive Statistics for Mexican Firms

Panel A: Regression Variables[a] (means (medians) reported in first (second) row)

Year	*R*	*PTE*	*Inv*	*AR*	*GM*	*S&A*	*Tax*	*Lev*
1993	**12.8%**	**25.6%**	−4.8%	**10.5%**	1.6%	0.6%	**−1.8%**	**40.6%**
n = 38	**7.1**	**12.5**	−3.1	**10.0**	−0.0	−0.6	**−0.2**	**41.1**
1994	6.2	0.7	**21.5**	**20.4**	2.2	0.6	**−1.9**	**47.1**
n = 53	0.3	4.6	**20.7**	**17.4**	1.4	1.7	−0.3	**46.2**
1995	−1.0	**13.1**	**−5.8**	**−5.6**	0.8	−2.1	**−1.7**	**45.7**
n = 70	−1.3	**9.5**	**−6.4**	**−11.4**	2.9	−1.3	−0.2	**47.4**
1996	2.0	**14.7**	**−5.6**	−4.5	**4.0**	**−4.6**	−0.8	**42.4**
n = 74	0.2	**11.9**	**−6.0**	**−7.6**	**1.5**	**−3.9**	−0.0	**43.2**
1997	−1.7	**11.8**	**8.0**	**6.8**	1.4	−2.1	**−1.8**	**41.9**
n = 91	−0.5	**8.1**	**8.0**	**5.3**	0.0	−1.4	**−0.3**	**43.5**
1998	**−20.0**	2.4	−0.8	0.4	0.1	**3.5**	**−1.4**	**43.3**
n = 118	**−16.6**	**3.5**	−0.1	−2.9	0.0	**3.4**	**−0.2**	**43.6**

Panel B: Correlation Analysis[a]

	PTE	*Inv*	*AR*	*GM*	*S&A*	*Tax*	*Lev*
R	**.39**	.05	.07	**−.14**	**−.30**	.05	**−.14**
PTE		−.04	.01	**−.14**	**−.17**	.01	**−.21**
Inv			**.30**	**.10**	**.10**	.01	.00
AR				−.07	**.14**	.01	−.04
GM					**−.15**	−.04	.08
S&A						.06	.01
Tax							**−.22**

Panel C: Properties of Analysts' Forecasts in Mexico[b] (means (medians) reported in first (second) row)

Year	*REV*	*FE*	Dispersion[c]	Analyst Following	No. of Forecast Revisions
1993	−0.68%	**−1.91%**	20.6%	9.9	1.6
n = 35	−0.37	**−1.04**	17.8	10.0	2.0
1994	**7.37**	**−9.18**	33.4	10.5	2.8
n = 46	**3.01**	**−4.25**	15.8	11.0	3.0
1995	**4.72**	−2.45	115.9	12.5	2.9
n = 59	**2.17**	0.67	41.8	11.0	2.0
1996	**−2.26**	0.15	367.7	13.0	3.2
n = 62	**−2.20**	1.10	28.6	12.5	3.0
1997	−1.30	−0.59	22.0	13.4	4.2
n = 72	−0.66	−0.57	22.6	13.0	3.0
1998	**3.08**	**−5.85**	19.1	8.5	3.2
n = 80	**1.52**	**−4.15**	16.8	8.0	2.0

(*continued on next page*)

TABLE 1 (continued)

---

[a] The sample consists of 444 firm-year observations from 1993–1998 with the requisite financial statement and price data from the Economatica database. To control for outliers, all variables in the model are winsorized at the 5 percent and 95 percent tails of their respective distributions. Numbers reported in **bold** indicate significance at the $\alpha = .05$ level (two-tailed) using parametric t-tests or the nonparametric Wilcoxon test for medians. *R* = market-adjusted 12-month return ending three months after the fiscal-year-end; *PTE* = pre-tax operating-earnings-per-share multiplied by 1 minus the tax rate in the preceding year, deflated by price at the beginning of the return accumulation period; *Inv* = percentage change in inventory minus the percentage change in sales; *AR* = percentage change in accounts receivable minus the percentage change in sales; *GM* = percentage change in sales minus the percentage change in gross margin; *S&A* = percentage change in selling and administrative expenses minus the percentage change in sales; *Tax* = pre-tax earnings-per-share multiplied by the difference in effective tax rates for the preceding year and the current year; and *Lev* = total liabilities divided by total assets.

[b] Analysts' forecasts were obtained from the I/B/E/S International Summary database. $REV = [(PostF_{t+1} - E_t) - (PreF_{t+1} - F_t)]/P_{t-1}$ where $PostF_{t+1}$ is the first available mean consensus forecast for earnings in year $t+1$ subsequent to the year t earnings announcement; E_t is realized earnings in year t; $PreF_{t+1}$ is the first available mean consensus forecast in year t for earnings in year $t+1$; F_t is the first available mean consensus forecast in year t for year t earnings; and P_{t-1} is price at the beginning of the returns accumulation period in year t. $FE = (E_t - F_t)/P_{t-1}$. *Dispersion* is measured as the standard deviation of all available forecasts that comprise F_t divided by F_t. *Analyst Following* is equal to the number of forecasts that comprise F_t. *No. of Forecast Revisions* represents the number of analysts who changed their forecasts of F_t either up or down from their last forecast.

[c] Dispersion was calculated for all companies where the analyst following is greater than 1. The n for each year for this variable is 1993 = 31; 1994 = 45; 1995 = 58; 1996 = 62; 1997 = 69; 1998 = 75.

comprise F_t, ranges from 8.5 in 1998 to 13.4 in 1997. The number of forecasts that changed either up or down from the previous period is shown in the last column of Panel C of Table 1. Revisions were not unusually frequent after the devaluation.

We also examined whether analysts' forecasts are updated on a timely basis. This information is of interest primarily because we later examine the effects of the fundamental signals on *REV* and want to know if the findings could be influenced by stale forecasts. In untabulated analyses, we found that, subsequent to the announcement of year t earnings, the first available forecast for year $t+1$ earnings ($PostF_{t+1}$) during non-devaluation years is available by March (April) of year $t+1$ for 94.8 percent (97.1 percent) of firms. The corresponding amounts after 1994 earnings are released are slower at 39.1 percent (100 percent), which means that $PostF_{t+1}$ (the forecast for 1995) was issued one month later for about half the firms (i.e., 94.8 percent less 39.1 percent).

This delay by analysts in providing $PostF_{t+1}$ might be due to the following reasons: Mexican companies were slow in reporting their earnings for 1994, and/or Mexican analysts required additional time in formulating a forecast of 1995 earnings after the announcement of 1994 earnings. We investigate these explanations by accessing the annual earnings announcement date from I/B/E/S and examining the timeliness of earnings reports across the sample period. For each year, we calculate the number of days from fiscal year-end to the earnings announcement date (not tabulated). Excluding 1994, the median ranges from 61 to 68 days. However, for 1994 the median value is 93 days. Thus, the delay by analysts in providing $PostF_{t+1}$ in 1994 is due to Mexican companies taking additional time to announce their 1994 earnings. While forecast staleness is not a serious problem, this delay brings up the question of whether our returns period for 1994 should be extended beyond March. As a result, we reran all analyses for 1994 using a return period ending with April 1995. None of our analyses changed substantively, so we present results using the March return period in order to be consistent with other years.

Association of Earnings and Fundamental Signals with Contemporaneous Returns

Table 2, Panel A, presents information about the association of annual earnings with returns for Model (1), which serves as a benchmark in subsequent tests of the incremental value relevance of the fundamental signals. As expected, earnings lose value relevance in 1994 due to the peso devaluation in the last week of the year. In contrast, when observations from the five non-devaluation years are pooled, earnings have considerable value relevance with an R^2 of 21.7 percent and the earnings coefficient of .694 has a t-value of 10.5. This result is robust after controlling for cross-sectional dependence, since the five-year mean of 0.596 is significant with an intertemporal t-statistic of 6.16. To provide a direct comparison between the devaluation and non-devaluation periods, we formally test whether the earnings coefficients differ between 1994 and the non-devaluation years. Pooling data for all six years and including a dummy variable for the 1994 observations, the t-value of 2.52 allows rejection of equality at the .05 significance level.

To determine whether the loss of value relevance in 1994 is due to economic conditions existing throughout the year or to the economic shock of the December devaluation, Table 2, Panel B, reports separate results for the first three quarters of 1994 and for the devaluation quarter. Since earnings become insignificant only in the devaluation quarter, we conclude that the loss of value relevance for 1994 annual earnings is due primarily to the economic shock of the devaluation.

Table 3, Panel A, examines whether the fundamental signals provide incremental explanatory power in comparison to annual earnings. For the 1994 devaluation year, the model containing both the fundamental signals and earnings has considerable explanatory power with an adjusted R^2 of 25.0 percent. This is substantially more than the adjusted R^2 of -0.2 percent for the earnings-only benchmark model, and the partial F-statistic of 3.94 shows that the improvement in explanatory power from including the six fundamental signals is statistically significant. The *S&A* signal is highly significant in 1994 ($t = -3.95$) and provides most of the improvement.

In the non-devaluation years, the fundamental signals again provide incremental explanatory power in comparison to annual earnings. The adjusted R^2 of 26.2 percent for the full model exceeds the adjusted R^2 of 21.7 percent for the benchmark earnings-only model, and the partial F-statistic of 4.95 indicates the difference is significant. The *GM* and *S&A* signals are statistically significant with t-statistics of -2.37 and -5.22, respectively. The corresponding intertemporal t-statistics are -2.55 and -1.96, and both signals have coefficients with the hypothesized negative sign in all five years.[10] In addition, the tests for a difference in coefficients indicate that the *PTE* and *S&A* coefficients for the devaluation year are significantly different than for non-devaluation years.

To determine whether the increased importance of the fundamental signals in 1994 is due to economic conditions prior to the devaluation or to the shock of the devaluation, Table 3, Panel B, reports separate results for the first three quarters of 1994 and for the fourth (devaluation) quarter. Examining the fourth quarter results, the adjusted R^2 of 24.7 percent exceeds the adjusted R^2 of -1.9 percent for the earning-only benchmark model (and is similar to the adjusted R^2 of 25 percent for the annual period). This increase in explanatory power is significant based on the partial F-statistic of 3.62. Also consistent with the annual results, the contribution of the fundamental signals is driven primarily by the *S&A* signal, which is significant with a t-value of -1.99. In contrast, the partial F-statistic for the first three quarters of 1994 shows that the fundamental signals as a group

---

[10] Note that, with a sample size of five years, the t-value of 1.96 for *S&A* is not statistically significant.

TABLE 2
Benchmark Model: Earnings/Return Relation for Sample of Mexican Firms[a]

$$R_{it} = \alpha_0 + \beta_1 E_{it} + \nu_{it}$$

Panel A: Tests for Annual Periods

	Intercept	E	Adj. R^2
Devaluation Year: 1994 (n = 53)			
Coefficient	0.065	0.213	−0.2%
(t-statistic)	(1.24)	(0.95)	
1993; 1995–1998 (n = 391)			
Pooled Coefficient	−0.115	**0.694**	21.7%
(t-statistic)	(−7.12)	**(10.5)**	
Mean Annual Coefficient	−0.085	**0.596**	
[No. Positive]	[1]	**[5]**	
(Inter-temporal t-statistic)	(−2.40)	**(6.16)**	
Difference in Earnings Coefficients			
Ho: $\beta_{1;94} - \beta_{1;93,95\text{–}98} = 0$			
t-statistic		**−2.52**	
p-value		**(.012)**	

Panel B: Tests For 1994 Subperiods

	Intercept	E	Adj. R^2
1st, 2nd, 3rd Quarters (n = 45)			
Coefficient	−0.030	**0.464**	7.2%
(t-statistic)	(−0.92)	**(2.10)**	
4th Quarter (n = 45)			
Coefficient	0.019	−0.109	−1.9%
(t-statistic)	(0.27)	(−0.45)	

---

[a] The sample consists of 444 firm-year observations from 1993–1998 with the requisite financial statement and price data from the Economatica database. To control for outliers, all variables in the model are winsorized at the 5 percent and 95 percent tails of their respective distributions. Numbers reported in **bold** indicate statistical significance at the $\alpha = .05$ level (one-tailed). R_{it} = market-adjusted 12-month return for firm i ending three months following the end of fiscal year t; and E_{it} = earnings-per-share for firm i in year t, deflated by price at the beginning of the return accumulation period.

do not have significant explanatory power—although the *S&A* signal is significant with a t-value of −1.67. Based on this result, we conclude that the increased importance of the fundamental signals in the annual analyses is due to the economic shock of the devaluation event, as opposed to economic conditions that triggered the devaluation.

Usefulness of Fundamental Signals in Predicting Future Earnings, Forecast Revisions, and Forecast Errors

Table 4 reports results from estimating Equation (3) with three different dependent variables: one-year-ahead earnings change (ΔE_{t+1}), analysts' forecast revisions (REV_t), and analysts' forecast errors (AFE_{t+1}). Examining the ΔE_{t+1} equation for both the devaluation year and the pooled non-devaluation years, the coefficient on *PTE* is significantly negative in both instances, which indicates reversion toward the mean. When revising their earnings forecasts (REV_t), analysts anticipate mean reversion, as indicated by a significant negative coefficient on *PTE* in both the devaluation and non-devaluation time periods. The amount

TABLE 3
Results of Earnings and Fundamental Analysis Signals Regressed on Contemporaneous Returns[a]

$$R_{it} = \alpha_0 + \beta_1 PTE_{it} + \beta_2 Inv_{it} + \beta_3 AR_{it} + \beta_4 GM_{it} + \beta_5 S\&A_{it} + \beta_6 Tax_{it} + \beta_7 Lev_{it} + \omega_{it}$$

Panel A: Tests for Annual Periods

	Intercept (+/−)	***PTE*** (+)	***Inv*** (+/−)	***AR*** (−)	***GM*** (−)	***S&A*** (−)	***Tax*** (+/−)	***Lev*** (−)	**Adj. R²**	**Benchmark Adj. R²**	**Partial F-statistic**
Devaluation Year: 1994 (n = 53)											
Coefficient	0.316	−0.164	−0.014	0.037	−0.423	**−0.985**	−0.608	−0.541	25.0%	−0.2%	**3.94**
(t-statistic)	(1.69)	(−0.63)	(−0.08)	(0.23)	(−1.23)	**(−3.95)**	(−0.61)	(−1.38)			
1993; 1995–1998 (n = 391)											
Pooled Coefficient	−0.064	**0.600**	0.069	0.016	**−0.287**	**−0.499**	0.556	−0.098	26.2%	21.7%	**4.95**
(t-statistic)	(−1.61)	**(8.77)**	(1.09)	(0.28)	**(−2.37)**	**(−5.22)**	(1.72)	(−1.18)			
Mean Annual Coefficient	−0.042	**0.553**	0.087	−0.057	**−0.253**	−0.423	1.241	−0.077			
[No. Positive]	[2]	**[5]**	[3]	[2]	**[0]**	**[0]**	[4]	[2]			
(Inter-temporal t-statistic)	(−0.57)	**(4.21)**	(1.01)	(−0.50)	**(−2.55)**	(−1.96)	(2.19)	(−0.61)			
Difference in Earnings Coefficients											
Ho: $\beta_{i;94} = \beta_{i;93,95-98}$											
t-statistic		**−3.21**			−0.42	**−2.03**					
p-value		**(.001)**			(0.675)	**(.043)**					

(*continued on next page*)

TABLE 3 (continued)

Panel B: Tests for 1994 Subperiods[b]

	Intercept (+/−)	*PTE* (+)	*Inv* (+/−)	*AR* (−)	*GM* (−)	*S&A* (−)	*Tax* (+/−)	*Lev* (−)	Adj. R^2	Benchmark Adj. R^2	Partial F-statistic
1st, 2nd, 3rd Quarters (n = 45)											
Coefficient	0.039	0.384	−0.124	0.072	−0.300	**−0.409**	0.179	−0.169	9.4%	7.2%	1.21
(t-statistic)	(0.42)	(1.59)	(−0.88)	(0.82)	(−0.75)	**(−1.67)**	(0.10)	(−0.81)			
4th Quarter (n = 45)											
Coefficient	0.162	−0.449	−0.077	0.204	0.547	**−1.420**	1.004	−0.600	24.7%	−1.9%	**3.62**
(t-statistic)	(0.72)	(−2.08)	(−0.49)	(1.83)	(0.47)	**(−1.99)**	(0.71)	(−1.38)			

---

[a] The sample consists of 444 firm-year observations from 1993–1998 with the requisite financial statement and price data from the Economatica database. To control for outliers, all variables in the model are winsorized at the 5 percent and 95 percent tails of their respective distributions. Numbers reported in **bold** indicate significance consistent with the direction prediction at the α = .05 level (one-tailed or two-tailed as indicated in the column heading). The partial F-statistic tests whether the fundamental variables provide a statistically significant explanatory power for returns over the benchmark earnings model.

R = market-adjusted 12-month return for firm *i* ending three months following the end of fiscal year *t*;
PTE = pre-tax operating-earnings-per-share multiplied by 1 minus the tax rate in the preceding year, deflated by price at the beginning of the return accumulation period;
Inv = percentage change in inventory minus the percentage change in sales;
AR = percentage change in accounts receivable minus the percentage change in sales;
GM = percentage change in sales minus the percentage change in gross margin;
S&A = percentage change in selling and administrative expenses minus the percentage change in sales;
Tax = pre-tax earnings-per-share multiplied by the difference in effective tax rates for the preceding year and the current year; and
Lev = total liabilities divided by total assets.

[b] The first subperiod uses deflated earnings for the first three quarters of the year regressed on market-adjusted returns beginning April 1, 1994 (so that calendar 1993 earnings have been released) and extending through November 30, 1994 (so that third quarter earnings have been released). The second subperiod uses fourth quarter earnings regressed on returns surrounding the devaluation event, covering December 1, 1994 through March 31, 1995 (see Figure 1, Panel B).

TABLE 4
Results of Earnings and Fundamental Analysis Signals Regressed on Future Earnings Changes, Analysts' Forecast Revisions, and Analysts' Forecast Errors

$$Y_{it} = \alpha_0 + \beta_1 PTE_{it} + \beta_2 Inv_{it} + \beta_3 AR_{it} + \beta_4 GM_{it} + \beta_5 S\&A_{it} + \beta_6 Tax_{it} + \beta_7 Lev_{it} + \omega_{it}$$

Panel A: Time Period

	Intercept (+/−)	*PTE* (−)	*Inv* (+/−)	*AR* (−)	*GM* (−)	*S&A* (−)	*Tax* (+/−)	*Lev* (−)	Adj. R^2
Devaluation Year: 1994 (n = 46)									
$Y = \Delta E_{t+1}$	0.137	**−0.508****	0.006	−0.030	**−0.282****	**−0.517****	0.193	−0.053	57.9%
$Y = REV_t$	−0.039	**−0.169***	−0.021	0.012	**−0.150***	**−0.140****	−0.361	0.241	51.4%
$Y = AFE_{t+1}$	0.120	−0.126	0.001	−0.035	**−0.225***	**−0.337****	0.085	−0.228	26.7%
1993, 1995–1998 (n = 308)									
$Y = \Delta E_{t+1}$									
Pooled Coefficient	0.039	**−0.172****	−0.057	0.025	**−0.307****	**−0.194****	**−0.467***	−0.053	9.9%
Mean Annual Coefficient	**0.057***	−0.088	0.007	0.042	**−0.251***	−0.179	−0.182	−0.146	
$Y = REV_t$									
Pooled Coefficient	−0.012	**−0.111****	−0.021	0.002	−0.057	0.032	**−0.320****	0.072	15.9%
Mean Annual Coefficient	−0.008	−0.078	−0.035	0.019	0.035	0.022	**−0.201***	0.042	
$Y = AFE_{t+1}$									
Pooled Coefficient	0.031	−0.025	0.001	0.019	**−0.266****	**−0.240****	0.002	**−0.168****	11.0%
Mean Annual Coefficient	0.043	0.037	0.078	0.028	**−0.267***	−0.198	0.085	−0.242	
Difference in Earnings Coefficients									
Ho: $\beta_{i;94} = \beta_{i;93,95\text{–}98}$									
$Y = \Delta E_{t+1}$	t-statistic	**−2.18***			0.14	**−2.49***	1.27	0.00	
$Y = REV_t$	t-statistic	−0.74			−1.05	**−2.65****	−0.16	1.55	
$Y = AFE_{t+1}$	t-statistic	−0.72			0.25	−0.82	0.18	−0.30	

Panel B: Dependent Variable

	Intercept (+/−)	*PTE* (−)	*Inv* (+/−)	*AR* (−)	*GM* (−)	*S&A* (−)	*Tax* (+/−)	*Lev* (−)	Adj. R^2
Forecast Revisions Surrounding December 1994 Devaluation for 1995 Annual Earnings	0.005 (0.45)	−0.040 (−1.54)	−0.005 (−0.48)	0.009 (0.95)	**−0.170** **(−4.38)**	**−0.050** **(−2.19)**	0.115 (1.25)	0.011 (0.48)	44.7%

(continued on next page)

TABLE 4 (continued)

---

*, ** Indicate significance at the $\alpha = .05$ and .01 levels, respectively, using a one-tailed test where directional predictions are provided; otherwise, two-tailed. All significant amounts are reported in **bold** type.

The sample consists of 354 firm-year observations from 1993–1998 with the requisite financial statement and price data from the Economatica database and analysts' forecast data from I/B/E/S. To control for outliers, all variables in the model are winsorized at the 5 percent and 95 percent tails of their respective distributions.

Panel A: ΔE_{t+1} = actual earnings-per-share in year $t+1$ minus actual earnings-per-share in year t, deflated by price at the beginning of the returns accumulation period (all data obtained from I/B/E/S);

REV_t = analyst earnings forecast revision for year $t+1$ using data obtained from I/B/E/S:

$$[(PostF_{t+1} - E_t) - (PreF_{t+1} - F_t)]/P_{t-1}$$

where:

$PostF_{t+1}$ = the first forecast of $t+1$ earnings available subsequent to the announcement of time period t earnings;
$PreF_{t+1}$ = the first forecast of $t+1$ earnings available in time period t;
F_t = the first forecast of t earnings available in time period t;
E_t = actual earnings for time period t;
P_{t-1} = price at the beginning of the returns accumulation period.

AFE_{it+1} = actual earnings-per-share in year $t+1$ minus the first forecasted earnings-per-share for year $t+1$ subsequent to year t's earnings announcement, deflated by price at the beginning of the returns accumulation period;
PTE = pre-tax operating-earnings-per-share multiplied by 1 minus the tax rate in the preceding year, deflated by price at the beginning of the return accumulation period;
Inv = percentage change in inventory minus the percentage change in sales;
AR = percentage change in accounts receivable minus the percentage change in sales;
GM = percentage change in sales minus the percentage change in gross margin;
$S\&A$ = percentage change in selling and administrative expenses minus the percentage change in sales;
Tax = pre-tax earnings-per-share multiplied by the difference in effective tax rates for the preceding year and the current year; and
Lev = total liabilities divided by total assets.

Panel B: The February 1995 mean consensus forecast minus the December 1994 mean consensus forecast for 1995 earnings deflated by price as of April 1, 1994. The fundamental signals are calculated using accounting balances reported in the third quarter of 1994. Coefficients reported with t-values in parentheses.

of the anticipation is sufficient so that *PTE* is not significant in explaining one-year-ahead forecast errors (AFE_{t+1}).[11]

With respect to the fundamental signals in the devaluation year, we find that the 1994 *GM* and *S&A* signals have a significantly negative association with the change in earnings, and analysts' revisions are significantly influenced by both signals. However, the predictive information provided by the signals is underutilized because both signals are significant in explaining 1995 forecast errors (AFE_{t+1}). This underutilization of information suggests that investors may be able to earn abnormal returns by trading on the information provided by these signals.

For the pooled non-devaluation years, we find that the *GM* and *S&A* signals again have a significant negative association with future earnings changes. Analysts do not use this information in their forecast revisions (whereas they used the information but not fully in the devaluation year). As a result, both signals are significantly associated with forecast errors. The underutilization of the information in those two signals indicates that investors may be able to earn abnormal returns even during non-devaluation years. Examining the other signals, the tax signal is also associated with future earnings changes, but analysts use this information in their forecast revisions, so it is not significant in explaining forecast errors. One other signal, leverage, is associated with forecast errors.

The bottom three rows of Table 4 test whether the coefficients differ significantly between devaluation and non-devaluation years on any of the variables found to be significant in explaining earnings changes, forecast revisions, or forecast errors. For $\Delta E_{t+1,}$ the tests show that *PTE* is significantly different with a greater extent of mean reversion in the devaluation year than in the non-devaluation years (t-value of −2.18). However, the effect of *PTE* on forecast revisions and forecast errors does not differ between the periods. Another variable with a significant difference is the *S&A* signal, which has a higher negative association with future earnings changes (ΔE_{t+1}) in the devaluation year than in non-devaluation years (t-value of −2.49). Analysts recognize this relationship in revising their forecasts, since their use of the *S&A* signal is significantly greater in the devaluation year compared to other years (t-value of −2.65). As a consequence, the forecast error is not significantly different between the periods. None of the other fundamental signals have a statistically different effect between the devaluation year and the non-devaluation years in predicting earnings changes, forecast revisions, or forecast errors.

The forecast revisions (REV_t) reported in Table 4, Panel A, use information about operations throughout 1994. To determine whether the revisions are due to the December devaluation, we next consider analysts' forecast revisions during the two months immediately surrounding the devaluation. Table 4, Panel B, compares the last monthly consensus forecast issued prior to the devaluation (in December 1994)[12] to the last consensus forecast issued before actual earnings are released (in February 1995).[13] We then regress fundamental signals measured using third quarter 1994 financial statement amounts on this forecast revision to determine whether the fundamental signals are associated with the information used by analysts. Corresponding to the results for REV_t reported on Panel A, analysts place significant weight on two components of operating performance, namely, a company's past ability to earn high profit margins (*GM*) and maintain control over operating expenses (*S&A*). Finding similar results indicates that the devaluation event, rather than economic conditions earlier during 1994, are driving the significance of the *GM* and *S&A* signals in REV_t.

---

[11] This discussion relies on the pooled data, although we also present mean annual coefficients for the non-devaluation years. The latter are often statistically insignificant because only five annual observations are available for the statistical tests.

[12] I/B/E/S indicates that the mean consensus forecasts are as of the Thursday before the third Friday of every month, which would be as of the 15th for the December 1994 forecast period—just four business days prior to the beginning of the devaluation.

Hedged Portfolio Trading Strategy

Our final analysis investigates the possibility that abnormal returns can be earned using a trading strategy that exploits information in the fundamental signals following the December 1994 devaluation. The trading strategy involves construction of hedged portfolios. The methodology has been previously used by Abarbanell and Bushee (1998) to investigate the profitability of trading on the LT signals in U.S. markets. This section provides only a brief description of the hedged portfolio methodology. For further details, the reader should refer to Abarbanell and Bushee (1998) and Fama and Macbeth (1973).

We add to our fundamental analysis model two risk proxies: firm size ($SIZE_t$), measured as the natural log of the market value of equity at the end of year *t*, and the market-to-book ratio (MB_t), calculated as the end-of-year market value of equity divided by the book value of equity. These control variables have been used in prior studies (e.g., Fama and French 1992; Fama 1998) to reduce the chance that any significant returns from the hedged portfolios are due to omitted risk factors.[14]

For every year, we rank each fundamental signal, earnings, and risk proxy into deciles (0 to 9) and divide each decile rank by 9 to create variables with a range from 0 to 1. Next, we change the sign of each variable so that its expected association with future returns is positive.[15] To determine the hedged portfolio return that would be realized by optimally weighting the portfolio based on the fundamental signals, we estimate the following equation:

$$R_{it+1} = \alpha_0 + \beta_1 PTE_{it} + \beta_2 SIZE_{it} + \beta_3 MB_{it} + \beta_4 Inv_{it} + \beta_5 AR_{it} + \beta_6 GM_{it} + \beta_7 S\&A_{it} + \beta_8 Tax_{it} + \beta_9 Lev_{it} + \omega_{it}, \tag{5}$$

where R_{it+1} is a market-adjusted return where the accumulation period begins in a month after the information in the fundamental signals is disclosed. The coefficients from Equation (5) represent the hedged abnormal return from a zero-investment portfolio that is optimally formed to exploit the information contained in that signal. The sum of β_4 through β_9 represents the net portfolio return from exploiting the information contained in all the signals, while controlling for year *t* earnings, size, and the market-to-book ratio.

---

[13] For every company in our sample, 1994 earnings had not yet been reported at the time the I/B/E/S February forecasts were published. Further, of the 40 companies with sufficient quarterly financial statement and analyst forecast data to be included in our tests, every one had a non-zero forecast revision, indicating the analysts were reacting to the devaluation. Also, 60 percent of the forecast revisions were negative.

[14] Abarbanell and Bushee (1998) use equity beta to control for risk. Many of the firms in our sample are thinly traded, which causes problems in measuring beta. Moreover, beta has been criticized as a poor measure of risk (Fama and French 1992).

[15] This step is executed by dividing all the fundamental signals by −9 except for *Inv*.

Table 5 presents the results from estimating Equation (5). To provide a benchmark, we also provide results from a trading strategy that is based on actual realized earnings in year $t+1$, which is entitled the "perfect foresight strategy," As expected, knowing future earnings levels results in a significantly positive abnormal return in most periods. The only exception is for 1993, which is consistent with the insignificant association between earnings and returns in 1994 (i.e., having perfect foresight of 1994 earnings in 1993 does not produce an abnormal return).[16]

The first set of results in Table 5 measure the abnormal returns during periods immediately following the devaluation.[17] We find that using a fundamental analysis strategy that exploits the information contained in the fundamental signals results in a significant abnormal return following the devaluation. Consistent with results in Table 4, the signals with significant coefficients in Equation (5) for this return window are the same that are significant in predicting future forecast errors, *GM* and *S&A*. Given that *GM* and *S&A* have been consistently the strongest signals throughout our analyses, the last columns present returns when the hedge portfolios are formed using only information in these two signals. The hedged returns are again highly significant in the extended windows following the devaluation. The last part of Table 5 investigates the success of the fundamental analysis strategy in the other years of our sample period. None of the fundamental analysis strategy returns are significant, and we do not find any individual fundamental signal coefficients to be significant.[18]

Although the method used to assess the hedged portfolio strategy in Table 5 is desirable in that it controls for risk factors while simultaneously considering the combined effect of all fundamental signals, we cannot determine how much of the abnormal returns are driven by the buy-side (long) portfolio versus the sell-side (short) portfolio. If the returns are driven by the sell-side portfolio, this would suggest that the strategy is less implementable since transaction costs would be more substantial. In Table 6, we present returns from an alternative hedged trading strategy. This strategy groups the sample firms by year into four portfolios based on individual fundamental signals that were derived from the 1994 annual report. The most (least) favorable change in the associated signal was labeled the long (short) portfolio. The mean and median returns to the portfolios assume that every security within each portfolio is equally weighted. Table 6 presents abnormal returns for the May 1, 1995, through April 30, 1996 return window for the *GM* and *S&A* signals. Consistent with the results in Table 5, statistical significance is found for each signal, and more importantly, we document significant buy-side returns for both signals.

To summarize, while the *GM* and *S&A* fundamental signals have a statistically significant association with contemporary returns, analysts underutilize the information in those signals, providing an opportunity for investors to earn abnormal returns after the 1994 devaluation. The market therefore does not fully impound the additional information provided by fundamental analysis after the economic shock.

---

[16] Size is significant in both the fundamental signal and perfect foresight equation for years 1994 and 1997. *MB* is significant in the fundamental signal (perfect foresight) equation in 1993 and 1994 (1997).

[17] Given the delay of companies reporting 1994 financial results, the return windows based on 1994 data are extended one month.

[18] This result differs from Abarbanell and Bushee (1998) who found an average 12-month cumulative size-adjusted abnormal return of 13.2 percent for a sample of U.S. firms over the period 1974–1988.

TABLE 5
Mean Hedged Portfolio Return from Fundamental Analysis and Perfect Foresight Strategies[a]

Fundamental Signal Equation:

$$R_{it+1} = \alpha_0 + \beta_1 PTE_{it} + \beta_2 SIZE_{it} + \beta_3 MB_{it} + \beta_4 Inv_{it} + \beta_5 AR_{it} + \beta_6 GM_{it} + \beta_7 S\&A_{it} + \beta_8 Tax_{it} + \beta_9 Lev_{it} + \omega_{it}$$

Perfect Foresight Equation:

$$R_{it+1} = \gamma_0 + \gamma_1 E_{it+1} + \gamma_2 SIZE_{it} + \gamma_3 MB_{it} + \varepsilon_{it}$$

			Perfect Foresight Stragtegy		Fundamental Analysis Strategy			Fundamental Analysis Strategy Using Only GM and S&A Signals	
Return Window	**Signals Derived From**	**n**	$\hat{\gamma}_1$	**p-value**	**Mean Hedged Return =** $\sum_{j=4}^{9} \hat{\beta}_j$	**p-value**	**Signals with Significant Coefficients**	**Mean Hedged Return =** $\sum_{j=6}^{7} \hat{\beta}_j$	**p-value**
Hedged Portfolio Returns Based on 1994 Data									
Mar. 1, 1995–Apr. 30, 1996	1994 3rd Q Data	44	0.805	<.01	0.973	.07	*GM, S&A*	0.907	<.01
May 1, 1995–Apr. 30, 1996	1994 Annual Data	51	0.928	<.01	0.585	.10	*GM, S&A*	0.904	<.01
Hedged Portfolio Returns Based on Non-1994 Data									
Apr. 1, 1994–Mar. 31, 1995	1993 Annual Data	38	0.202	.18	−0.241	.62	*	0.099	.76
Apr. 1, 1996–Mar. 31, 1997	1995 Annual Data	69	0.483	<.01	0.211	.42	*	−0.060	.71
Apr. 1, 1997–Mar. 31, 1998	1996 Annual Data	74	0.382	.01	0.112	.76	*	0.274	.13
Apr. 1, 1998–Mar. 31, 1999	1997 Annual Data	89	0.465	<.01	0.131	.62	*	−0.121	.42
Apr. 1, 1999–Mar. 31, 2000	1998 Annual Data	97	0.447	<.01	−0.128	.42	*	−0.040	.85

(continued on next page)

TABLE 5 (continued)

---

[a] The sample consists of all available observations from 1993–1998 with the requisite financial statement and price data from the Economatica database. To control for outliers, all variables in the model are winsorized at the 5 percent and 95 percent tails of their respective distributions.

R_{it+1} = future (relative to when the fundamental signals are calculated) market-adjusted return corresponding to the return window as indicated in the table;
PTE = pre-tax operating-earnings-per-share multiplied by 1 minus the tax rate in the preceding year, deflated by price at the beginning of the return accumulation period;
Inv = percentage change in inventory minus the percentage change in sales;
AR = percentage change in accounts receivable minus the percentage change in sales;
GM = percentage change in sales minus the percentage change in gross margin;
S&A = percentage change in selling and administrative expenses minus the percentage change in sales;
Tax = pre-tax earnings-per-share multiplied by the difference in effective tax rates for the preceding year and the current year;
Lev = total liabilities divided by total assets;
SIZE = the natural log of the market value of equity;
MB = market value of equity divided by the book value of equity; and
E = net income, deflated by price at the beginning of the return accumulation period.

TABLE 6
Portfolio Returns to Long and Short Positions in Mexican Companies Based on Individual Fundamental Signals[a]

	Funamental Signal = *GM*						Funamental Signal = *S&A*					
	Long Portfolio			Short Portfolio			Long Portfolio			Short Portfolio		
Year	n	Mean	Median	n	Mean	Median	n	Mean	Median	n	Mean	Median
1994	12	0.215*	0.190*	13	−0.198	−0.363	12	0.246*	0.221*	13	−0.327	−0.409

---

* The long (short) portfolio return is significantly greater than the short (long) portfolio return at the $\alpha = .01$ level using a two-tailed test.

[a] Each fundamental signal is defined as before. The entire sample is ranked into four portfolios based on the associated fundamental signal. The long (short) portfolio is comprised of firms with the most (least) favorable signal. The returns period runs from May 1, 1995 to April 30, 1996.

Supplemental Robustness Tests

Adding a Signal for Negative Earnings

Research in the U.S. shows that the relationship between earnings and returns is attenuated when earnings are negative (Hayn 1995). For 1994, we find that 39.6 percent of our sample firms have negative earnings (*E*) and 52.8 percent have negative earnings using the prior year's tax rate (*PTE*). To investigate the impact on our results of controlling for negative earnings, we set a dummy variable equal to 1 when earnings are negative (0 otherwise) and include it in our benchmark regression, Model (1), along with its interaction with earnings. The dummy variable main effect is insignificant in every year. However, the interaction with earnings (*E*) is significantly negative in 1994 and 1998 at the .05 level using a one-tailed test, and *E* becomes significantly positive in 1994.[19] The interaction term is insignificant over the 1993 to 1998 period in the cross-temporal means test.

Because research indicates that the association of earnings with returns is reduced by negative earnings, we investigate whether including an additional fundamental variable to indicate the sign of earnings could further enhance our ability to explain contemporaneous returns. This possibility is investigated by including a dummy variable (negative earnings equal to 1) as both an intercept and an interaction term by multiplying it times *PTE* in Equation (2). The primary results of our analyses are qualitatively unchanged, and inclusion of the dummy interaction term does not improve upon the hedged portfolio returns.

Constant Sample over Time

Data availability in Economatica allows us to increase the number of firms in our sample over time. We examine how our results are affected by retaining only those companies that have the requisite data for the entire sample period. Although the results with this smaller sample of 36 firms are somewhat weaker across years, the overall findings are generally consistent with the results reported in the tables. We still find that *GM* and *S&A* are consistently associated with contemporaneous returns (especially in the devaluation year where the adjusted R^2 is -2.2 percent for the benchmark model versus 30.4 percent for the fundamental signals model) and analysts tend to underutilize the information provided by these signals. When only *GM* and *S&A* are included in the model, the hedged portfolio provides a significantly positive abnormal return for 1994. Using the full model, however, the hedged return is not significant.

VI. CONCLUSION

Investors' need for forward-looking accounting information is greatly increased when an economic shock occurs during a reporting period. This is particularly true when the shock occurs late in the reporting period, so that current earnings information cannot be extrapolated to the future. Under these conditions, earnings lose value relevance and investors are forced to rely on other information. Currency devaluations provide a rich, natural experimental setting in which to investigate the effects of severe economic change on the usefulness of accounting information. The primary insight from research on the effects of a currency devaluation on accounting information is that earnings lose value relevance (Graham et al. 2000; Ho et al. 2001). This paper differs from prior research by investigating whether the detailed performance information provided in financial statements retains value relevance after a currency devaluation. This is accomplished using fundamental analysis, relying primarily on the signals developed by Lev and Thiagarajan (1993) (LT). More specifically, we test whether various fundamental signals (e.g., changes in gross margin or in selling and administrative expenses) provide forward-looking information to investors after the 1994 peso devaluation in Mexico.

---

[19] Graham et al. (2000) and Davis-Friday and Gordon (2002) perform similar analyses and also find that, after controlling for negative earnings in the devaluation year, positive earnings has a significant association with firm value.

For a sample of companies traded on the Mexican Bolsa, we find that earnings are not value relevant immediately following the 1994 currency devaluation. In contrast, the fundamental signals, especially the selling and administrative expense signal, provide value relevant accounting information (R^2 of 25 percent). Additional tests using interim accounting disclosures show that earnings remain value relevant during the first three quarters of 1994, but in the fourth (devaluation) quarter, earnings lose value relevance. In contrast, fundamental analysis, primarily the selling and administrative expense signal, results in significant value relevance in the devaluation quarter (R^2 of 24.7 percent). This finding indicates the increased contribution of fundamental analysis is due to the devaluation event, rather than to a gradual decline in economic conditions throughout 1994.

We also provide analyses that do not rely on an assumption that contemporaneous market reactions capture the full extent to which accounting information may be of value to investors. The primary purpose of those analyses is to determine (1) whether the fundamental signals provide forward-looking information after the 1994 devaluation and (2) whether analysts fully capture this information. To answer those questions, we examine associations of the fundamental signals with one-year-ahead earnings changes, analysts' revisions of their earnings forecasts, and analyst forecast errors. Two fundamental signals, gross margin and the selling and administrative expense signals, are significant in several of those analyses. Since the evidence indicates that analysts underutilize the information in 1994, an opportunity may exist in 1994 to use those signals to profit using a zero-investment trading strategy. Our tests show that a substantial profit can be earned in this manner. This test indicates that the market does not fully understand the increased value of fundamental analysis after an economic shock.

Our overall conclusion is that fundamental analysis assumes additional importance in the wake of the economic shock from a devaluation, but the market does not fully impound this fact. This finding presents the intriguing possibility that the more detailed accounting information used in fundamental analysis may provide value relevant information in the aftermath of other types of economic shocks (e.g., a terrorist attack, an oil embargo, or a labor strike).

REFERENCES

Abarbanell, J. S., and B. J. Bushee. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research* 35 (Spring): 1–24.

———, and ———. 1998. Abnormal returns to a fundamental analysis strategy. *The Accounting Review* 73 (January): 19–45.

Amir E., B. Lev, and T. Sougiannis. 1999. What value analysts? Working paper, New York University, New York, NY.

Anderson, M. C., D. Banker, and S. Janakiraman. 2000. Are selling, general, and administrative costs sticky? *Journal of Accounting Research* 41 (March): 47–63.

Brown, P., G. Foster, and E. Noreen. 1985. *Security Analyst Multi-Year Earnings Forecasts and the Capital Market.* Sarasota, FL: American Accounting Association.

Capstaff, J., K. Paudyal, and W. Rees. 1998. Analysts' forecasts of German firms' earnings: A comparative analysis. *Journal of International Financial Management and Accounting* 9: 83–116.

Davis-Friday, P., and E. Gordon. 2002. The effect of macroeconomic changes on the value-relevance of accounting information: The case of Mexico and the 1995 financial crisis. Working paper, University of Notre Dame, Notre Dame, IN.

Fama, E., and J. Macbeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (May-June): 607–636.

———, and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (June): 427–465.

———. 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49 (September): 283–306.

Francis, J., and D. Philbrick. 1993. Analysts' decisions as products of multi-task environment. *Journal of Accounting Research* 31 (Autumn): 216–230.

Gordon, E. A. 2001. Accounting for changing prices: The value relevance of historical cost, price level, and replacement cost accounting in Mexico. *Journal of Accounting Research* 39 (June): 177–200.

Graham, R., R. King, and J. Bailes. 2000. The value relevance of accounting information during a financial crisis: Thailand and the 1997 decline in the value of the baht. *Journal of International Financial Management and Accounting* 11 (Summer): 84–107.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (September): 125–153.

Ho, L. J., C. Liu, and P. S. Sohn. 2001. The value relevance of accounting information around the 1997 Asian financial crisis—The case of South Korea. *Asia-Pacific Journal of Accounting & Economics* 8: 83–107.

Instituto Mexicano de Contadores Publicos (IMCP). 1984. *Principios de Contabilidad Generalmente Aceptados*. Bulletin B-10, Reconocimiento de los Efectos de la Inflacion en la Informacion Financiera. Mexico City, Mexico: IMCP.

Lev, B., and S. R. Thiagarajan. 1993. Fundamental information analysis. *Journal of Accounting Research* 31 (Autumn): 190–215.

O'Brien, P. C. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10: 53–83.

Penman, S. H. 2001. *Financial Statement Analysis & Security Valuation*. New York, NY: McGraw-Hill/Irwin.

Editorial Data

The following table contains information about turnaround time for manuscripts (including revisions) on which editorial decisions were made in the 12-month period ended February 28, 2003. Turnaround time is the number of days between the date that the manuscript was received *and the date of the editor's letter to the author(s)*:

	Number of Manuscripts	Cumulative Number	Cumulative Percent
0 ≤ Days ≤ 30	70	70	18.45
31 ≤ Days ≤ 60	137	207	54.62
61 ≤ Days ≤ 90	121	328	86.54
91 ≤ Days ≤ 120	41	369	97.36
121 ≤ Days	10	379	100.00

The mean review time was 56.8 days; the median review time was 57 days.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal....No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of *The Chicago Manual of Style* (14th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be typed in 12-point font on one side of 8½ × 11" good quality paper and be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years.* All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form.* See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings.

Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix a, b, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says...."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by *The Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Date of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities.* 2nd edition. Homewood, IL: Dow Jones-Irwin.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge, U.K.: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–636.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, The University of Texas at Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting.* Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan or the like) should be submitted. Information that might identify the author(s) must be deleted from the instrument.
3. MS should be submitted via email as Microsoft® Word or PDF files to the *TAR* office at taruw@u.washington.edu. A submission letter, a separate cover page, and the MS containing no author identification should be emailed as separate files. The submission fee should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If unable to pay online, a check in U.S. funds, made payable to the American Accounting Association, may be sent to Professor Terry Shevlin, Senior Editor *TAR*, School of Business, Box 353200, University of Washington, Seattle, WA 98195-3200. The submission fee in U.S. funds is $75.00 for members and $100.00 for nonmembers of the AAA. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review.* If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting. "An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

JULY 2003 PLACEMENT ADS

*The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The **Review** is published on the first day of January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding.*

301 UNIVERSITY OF NOTRE DAME, Department of Accountancy, invites applications for position openings at the Assistant and Associate Professor levels. All research areas will be considered. Applicants for Assistant Professor positions should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong commitment to scholarly research and excellence in teaching. Salary is very competitive and research support is excellent. Please send *curriculum vitae* and an example of scholarly work to Professor Tom Schaefer, Chair, Department of Accountancy, Room 102D, Mendoza College of Business Administration, University of Notre Dame, Notre Dame, IN 46556-5646. The University of Notre Dame is an Equal Opportunity, Affirmative Action Employer.

302 CALIFORNIA STATE UNIVERSITY, LOS ANGELES, invites applications for positions in Accounting Information Systems (AIS), Financial, Cost/Managerial, Governmental, International, or Forensic Accounting. Start date is September 1, 2003. Rank will be commensurate with experience. Applicants must have a doctoral degree from an accredited institution of higher education in Accounting, AIS, or a closely related field. ABDs will be considered with doctorate completed prior to start of appointment. We are looking for faculty members who can provide a blend of teaching, scholarly activities, and service in a diverse urban community. Review of applications is ongoing and will continue until positions are filled. Write to Ms. Susan Bergstrom, Assistant to the Dean, College of Business and Economics, California State University, Los Angeles, 5151 State University Drive, Los Angeles, CA 90032-8120; or telephone (323) 343-2830. California State University, Los Angeles is an Equal Opportunity/Title IX/ADA Employer.

303 CASE WESTERN RESERVE UNIVERSITY, Weatherhead School of Management, invites applications for Assistant Professor of Accounting. Required qualifications include a doctorate in Accounting and a strong record of potential scholarship. A proven record of quality teaching is very important. Although candidates with a specialization in any Accounting area are invited to apply, we have a preference to those in the Financial and Managerial areas. CWRU currently offers undergraduate and master's level degree programs including a B.S. in Accounting and a Master's of Accountancy. In addition, faculty are heavily involved in M.B.A. program teaching and the supervision of doctoral students. All programs are AACSB-accredited. Additional information can be found at http://weatherhead.cwru.edu. Send *curriculum vitae* and other materials to Tim Fogarty, Department of Accountancy, Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106-7235; Email: tjf@cwru.edu. CWRU is an Affirmative Action/Equal Opportunity Employer. Applications from women and minority candidates are especially welcome.

304 CHEUNG KONG GRADUATE SCHOOL OF BUSINESS, Beijing, China, invites applications for a tenure-track position in the field of Accounting. The position can be at any level, but priority will be given to the level of Associate or Full Professor. Appointment will be from September 2003 or sooner. Applicant must have a Ph.D. and potential for excellence in research and teaching. For a senior appointment, applicant must have a demonstrated exceptional research and/or teaching record. We offer a competitive compensation package. Applicants should submit their *curriculum vitae,* any evidence supporting their teaching and research records, and ensure that three letters of recommendation are sent to Academic Associate Dean, Cheung Kong Graduate School of Business, Oriental Plaza, Tower E3, 1 East Chang An Avenue, Beijing 100738 P.R. China; Email: Recruit@ckgsb.com.

305 BAYLOR UNIVERSITY, Department of Accounting and Business Law, invites applications for an anticipated tenure-track position as Assistant Professor in Accounting Information Systems beginning Fall 2004. A Ph.D. (or ABD) in Accounting is required; professional certification is preferred. Applicants must demonstrate research and teaching excellence. Applications, including resume, three letters of reference, working paper or defended dissertation proposal, and evidence of an active Christian faith commitment should be sent to Dr. Charles E. Davis, Department of Accounting and Business Law, Baylor University, Box 98002, Waco, TX 76798-8002; Email: Charles_Davis@baylor.edu; Phone: (254) 710-6138. Applications will be reviewed beginning June 15, 2003 and will be accepted until the position is filled. To ensure full consideration, applications must be completed by September 15, 2003. Baylor University is a Baptist university affiliated with the Baptist General Convention of Texas. As an Affirmative Action/Equal Employment Opportunity Employer, Baylor University encourages minorities, women, veterans, and persons with disabilities to apply.

306 UNIVERSITY OF NEW ORLEANS invites applications for an Assistant Professor position for Fall 2003. Preference will be given to those with a primary teaching/research interest in Cost/Managerial or Financial/Governmental, with a secondary interest in Accounting Information Systems or Auditing. Candidates must have the Ph.D. with an Accounting major (or ABD with completion of the Ph.D. by the end of 2003). Salary is negotiable and competitive. The Department offers the undergraduate Accounting degree (over 550 majors) and two Master's degrees (including a Tax Option). The Department has separate AACSB accounting accreditation at both the graduate and undergraduate levels. Candidates should send a resume to Professor Philip J. Harmelink, Chairman, Department of Accounting, College of Business Administration, University of New Orleans, New Orleans, LA 70148; Email: pharmeli@uno.edu. The University of New Orleans is a member of the Louisiana State University System and an Affirmative Action, Equal Opportunity Employer.

307 OREGON STATE UNIVERSITY invites applicants for a tenure-track position in Accounting effective September 2004. Candidates should be able to teach in two of the following three areas: Financial, Auditing, or Information Systems. Responsibilities include research, quality teaching, and professional activities in the field of Accounting. Candidates should hold the appropriate doctorate. Professional certification is desirable. Salary is competitive; fringe benefits are excellent. Oregon State University is located in Corvallis, rated the second-best small city in America. Interested applicants should submit a letter of interest, current *curriculum vitae,* recent teaching evaluations, evidence of scholarly work, and a list of three references to Dr. Roger Graham, College of Business, 200 Bexell Hall, Oregon State University, Corvallis, Oregon 97331-2603. For full consideration, apply by September 30, 2003. Applications will be accepted until the positions are filled. Our faculty will meet informally with select candidates at the AAA Annual Meeting in Honolulu. OSU is an AA/EEO Employer.

308 PEPPERDINE UNIVERSITY, Graziadio School of Business and Management, is looking for an experienced Assistant or Associate Professor in Accounting for Fall 2004 who has established a solid research record in his or her field of expertise. We have both a full-time M.B.A. program and the largest part-time M.B.A. program in Southern California. This part-time program is taught in the evenings at six locations throughout the L.A. area. We value excellence in the classroom, recognize all types of scholarship, and encourage consulting activities. If you are ready to make a change, or thinking about making a change, we'd like to talk with you and tell you more about the Graziadio School and the uniqueness of Pepperdine University. Professor Richard Powell will be at the AAA Annual Meeting in Honolulu, or you can contact Professor Charles Hunt, Chair, Department of Accounting, Law, Finance, and Information Systems, Graziadio School of Business and Management, Pepperdine University, 400 Corporate Pointe, Culver City, CA 90230; Phone: (310) 568-5539; Email: charles.hunt@pepperdine.edu.

309 WESTERN NEW ENGLAND COLLEGE, School of Business, invites applications for a tenure-track Assistant/Associate Professor position in Accounting to begin in September 2003. Candidates are expected to hold a doctorate degree in Accounting or be in the advanced stages of completing their dissertation by the starting date. The successful candidate will demonstrate a strong commitment to both teaching in the undergraduate and graduate programs and to scholarly research. Candidates in all Accounting disciplines will be considered. Western New England College is a private, independent, coeducational institution founded in 1919. For more information about the College, visit our website at http://www.wnec.edu. Interested applicants should send information to Dr. Stanley Kowalski, Jr., Dean, School of Business, Western New England College, 1215 Wilbraham Road, Springfield, MA 01119. All applications should include a *curriculum vitae*, recent teaching evaluations, a recently completed research paper, and the names, addresses and telephone numbers of three references. Western New England College is an Equal Opportunity Employer.

310 UNIVERSITY OF MARYLAND, College Park, invites applications for a tenure-track/tenured position in Financial Accounting to start in August 2004. A strong research record is required. Send *curriculum vitae*, three letters of recommendation, and evidence of research to Professor Oliver Kim, The Robert H. Smith School of Business, University of Maryland, College Park, MD 20742-1815. The University of Maryland is an Equal Opportunity/Affirmative Action Employer. Women and minorities are especially encouraged to apply.

311 UNIVERSITY OF MARYLAND seeks applicants for a continuing nontenure-track teaching position in Accounting. The position is for nine months beginning August 2004, and is renewable subject to performance. Off-term teaching is available. Candidates must possess a doctorate and an excellent teaching record, especially at the Master's level. Teaching breadth will be valued. Certain student support and service responsibilities will be expected. Send a *curriculum vitae*, three letters of recommendation, and evidence of teaching effectiveness to Professor Oliver Kim, The Robert H. Smith School of Business, University of Maryland, College Park, MD 20742-1815. The University of Maryland is an Equal Opportunity/Affirmative Action Employer. Women and minorities are especially encouraged to apply.

312 SUNY COLLEGE AT BROCKPORT seeks applicants for a tenure-track Assistant Professor of Accounting position in the Department of Business Administration and Economics, starting in Fall 2003. Candidates are required to have a Ph.D. or D.B.A. in Accounting. Scholarly potential demonstrated by publications or high-quality work in progress and research plan is needed. SUNY Brockport offers undergraduate programs in Accounting, Business Administration, and International Business. All programs are accredited by AACSB International. The College is located in a large, dynamic business community, near the Lake Ontario shoreline. See our website for full vacancy announcement, http://www.brockport.edu/~hr/facultyandlibrarian.htm. Send letter of application, *curriculum vitae*, teaching evaluations, graduate transcripts, and three references to Ms. Adrienne Collier, Affirmative Action Officer, 421 Allen Administration Building, 350 New Campus Drive, SUNY College at Brockport, Brockport, NY 14420-2929. Informal inquiries may be addressed to Dr. Sandeep Singh; Phone: (585) 395-5519 or Email: ssingh@brockport.edu.

313 LOUISIANA STATE UNIVERSITY, Department of Accounting in the E. J. Ourso College of Business Administration, invites applications for one or more anticipated tenure-track openings to begin August 2004. Applications at the Assistant and Associate Professor rank will be considered in all disciplines, with priority given to Cost/Managerial or Financial Accounting specialties. Required qualifications include: (Associate Professor) Doctoral degree in Accounting; (Assistant Professor) Doctoral degree in Accounting be completed, or near completion, by the date of appointment (if degree is not completed by date of appointment, rank will initially be at the Instructor level). Additional desired qualifications include: CPA responsibilities; teaching undergraduate and graduate courses; active involvement in academic research; providing services to the profession and the university. Applications will be accepted until December 1, 2003, and will continue until candidates are selected. Please send a letter of application, *curriculum vitae* (including email address), and references to Chair, Department of Accounting, 3101 CEBA Building, Louisiana State University, Ref: Log #0900, Baton Rouge, LA 70803. LSU is an Equal Opportunity/Equal Access Employer.

314 UNIVERSITY OF MICHIGAN–FLINT seeks highly qualified applicants for a tenure-track position in Accounting beginning Fall 2003. Candidates should possess an earned doctorate in Accounting or closely related area. Candidate with record of successful teaching and research experience is preferred. A teaching background in the areas of AIS/MIS or Taxation is desirable. Salary and rank depends upon qualifications. UM–Flint's Business programs are accredited by AACSB International. Applicants should submit a resume, sample of professional writing, course evaluations, names, addresses, and phone numbers of at least three references. Contact Dr. Fred E. Williams, School of Management, University of Michigan–Flint, 303 E. Kearsley, Flint, MI 48502-1950; Phone: (810) 762-3164; Email: tedwill@umflint.edu. The University of Michigan–Flint is an Equal Opportunity/Affirmative Action Institution.

315 SEATTLE UNIVERSITY invites applications for a tenure-track Accounting position at the Assistant or Associate rank starting Fall Quarter 2004. Applicants for the position should possess a Ph.D. The candidate is expected to have an established record as an effective teacher and a demonstrated capacity to conduct publishable research. Candidates should have a teaching interest in Financial Accounting, including Advanced and Auditing. Seattle University's business school is AACSB-accredited with approximately 1,400 students in its graduate and undergraduate programs. A Master of Professional Accounting and M.B.A. Accounting concentration are offered. Seattle University, a Catholic and Jesuit urban university, is committed to the centrality of teaching, learning, and education of the whole person. Send inquires and resumes by October 1, 2003 to Susan Weihrich, Chair, Department of Accounting, Albers School of Business & Economics, Seattle University, 900 Broadway, Seattle, Washington 98122: Phone: (206) 296-5784; Email: weihrich@seattleu.edu. Seattle University is an Equal Employment Opportunity/Affirmative Action Employer.

316 UNIVERSITY OF NORTH ALABAMA is accepting applications for a tenure-track position in Accounting. Rank will depend upon qualifications. An earned doctorate in Accounting (or ABD) is preferred. Candidates from all Accounting disciplines may apply, but Tax or Accounting Information Systems candidates are preferable. Professional certification, experience in Accounting, and experience in teaching are highly desirable. Responsibilities include undergraduate and graduate teaching, academic advisement, scholarly research, and university service. Involvement in professional Accounting activities is encouraged. Review of applications will begin immediately. Applications will be accepted until the position is filled. See UNA website at http://www.una.edu/. Submit letter of application, *curriculum vitae,* unofficial copies of transcripts, and names, addresses, and telephone numbers of three references to Director of Human Resources and Affirmative Action, University of North Alabama, UNA Box 5043, Florence, AL 35632-0001. UNA is an Equal Opportunity Employer committed to achieving excellence and strength through diversity.

317 SEOUL NATIONAL UNIVERSITY, College of Business Administration, is seeking candidates with a Ph.D. and established research publication record to fill one professorship position at all levels in Accounting. The appointment is for 1–3 year term, with possible conversion to a tenure-track appointment. The successful candidate will be expected to teach at both the graduate and the undergraduate levels, and to supervise doctoral thesis research. The expected starting date is September 1, 2003, but can be deferred to 2004. Interested individuals are encouraged to submit their *curriculum vitae*, research papers, and the contact details of three references to Associate Dean Chang-woo Lee, College of Business Administration, Seoul National University, Kwanak-ku, Seoul, Korea 151-742; Phone: (822) 880-6901; Email: changwoo@snu.ac.kr. For more information, please visit http://cba.snu.ac.kr.

318 GRAND VALLEY STATE UNIVERSITY, Seidman School of Business, invites applications for Assistant/Associate Professor in Accounting. Tenure-track position is available Fall 2004 and open until filled. Need Ph.D. or D.B.A. (by December 2004) in Accounting or related field. Professional certification is desirable. Cost/Managerial Accounting and Financial Accounting are primary needs; other fields may be considered. Applicant must demonstrate the ability to teach effectively and conduct research. The Department and School are both accredited by the AACSB. Salary is competitive. For consideration, send resume, three references, and sample of teaching evaluations to Ms. Kathy Goralski, Accounting Search Committee, Department of Accounting and Taxation, Grand Valley State University, 401 West Fulton Street, Grand Rapids, MI 49504. To interview at the AAA Annual Meeting in Honolulu, please contact Professor Parvez Sopariwala, Accounting Department, Phone: (616) 331-7406; or Email: sopariwp@gvsu.edu by July 30, 2003. Position is open until filled. Grand Valley State University is an EO/AA Institution.

319 UNIVERSITY OF TENNESSEE, KNOXVILLE, is accepting applications for two full-time, tenure track positions (Assistant or Associate Professor) to begin January 1, 2004 or August 1, 2004. The second position is contingent on funding authorization. The positions are for Information Management within the Department of Accounting. Accordingly, candidates should hold a doctoral degree with a specialization in Accounting Information Systems or Management Information Systems, or expect to complete all degree requirements by his or her starting date. Although Information Systems is preferred, we also invite applications in other areas. Review of applications will continue until positions are filled. Send applications to Dr. Kenneth E. Anderson, Search Committee Chair, Department of Accounting and Business Law, CBA, University of Tennessee, Knoxville, Tennessee 37996-0560; Email: kea@utk.edu. UT Knoxville is an EEO/AA/Title VI /Title IX/Section 504/ADA/ADEA Institution in the provision of its education and employment programs and services. Applications from minorities and women are strongly encouraged.

320 BOISE STATE UNIVERSITY, Department of Accountancy, invites applications for a tenure-track position in Taxation at the Assistant Rank starting Fall 2004, subject to budgetary approval. A Ph.D. in Accounting is required (ABDs will be considered). Professional experience and certification are desired. Expectations include a commitment to high-quality teaching, scholarship leading to refereed publication, and university and community service. Both the College and Department's undergraduate and graduate programs are AACSB-accredited. Salary is competitive in the western region. Candidates should send a letter of interest, resume, and three letters of reference to Dr. Paul Bahnson, College of Business and Economics, Boise State University, 1910 University Drive, Boise, ID, 83725-1610. Faculty will be available to meet with applicants at the AAA Annual Meeting in Honolulu. The closing date for applications is November 1, 2003. Boise State University is an EEO/AA Employer.

321 UNIVERSITY OF HAWAII AT MANOA, College of Business Administration, School of Accountancy, invites applications for Assistant Professor or Associate Professor, Position Number 0085163. This general-funds, tenure-track position is full-time and begins August 1, 2003. Duties include teaching and research in the area of Accounting Information Systems, Auditing, or Financial Accounting, with interests in other Accounting subjects and performance of related tasks as assigned. Minimum qualifications for Assistant Professor: Ph.D., D.B.A., or ABD in Accounting; demonstrated ability to carry out and publish research; and demonstrated ability in teaching at the university level. Minimum qualifications for Associate Professor: Same as Assistant Professor; in addition, a minimum of four years of full-time college/university teaching at the rank of Assistant Professor or equivalent. Desirable qualifications include CPA, CMA, professional Accounting experience. Salary is negotiable based on qualifications and experience. Submit letter of application (note position number), three letters of recommendation, and an updated resume to Dr. David Yang, Chair of Recruiting Committee, School of Accountancy, College of Business Administration, University of Hawaii at Manoa, 2404 Maile Way, Honolulu, Hawaii 96822. Recruitment will continue until position is filled; first screening date is May 1, 2003. University of Hawaii at Manoa is an Equal Opportunity, Affirmative Action Institution.

322 SAINT LOUIS UNIVERSITY, a Catholic Jesuit institution dedicated to education, research, and healthcare, invites applications for two tenure-track positions in Accounting at the Assistant/Associate/Professor level beginning Fall 2004. Duties include commitment to teaching excellence; quality research and publication; service to the Department and University. All areas of Accounting will be considered. Ph.D. or near completion is required. Department representatives will attend the AAA Annual Meeting in Hawaii. Candidates are strongly encouraged to schedule an interview by contacting Dr. Ananth Seetharaman, Department of Accounting, Saint Louis University, 3674 Lindell Blvd., St. Louis, MO 63108; Phone: (314) 977-3828; Fax: (314) 977-1473; Email: acas@slu.edu. The positions are subject to budget approval. Saint Louis University is an Affirmative Action, Equal Opportunity Employer and encourages nominations and applications of women and minorities.

323 VILLANOVA UNIVERSITY invites applications at all levels for a tenure-track position in Auditing. The candidate should be capable of high-quality research and teaching at both the undergraduate and graduate levels, and is expected to provide leadership in developing/teaching contemporary auditing courses that are driven by problem-based learning. The successful candidate's record must reflect continued professional development and interaction with Accounting and Auditing firms. A doctorate in Accounting is required and CPA certification is preferred. Salary is competitive and commensurate with rank and carries with it the possibility of a professorship. The College of Commerce and Finance and its accounting program are accredited by the AACSB. Representatives will be attending the AAA Annual Meeting in Hawaii. Interested candidates should send their *curriculum vitae* to Anthony H. Catanach, Jr., College of Commerce & Finance, Villanova University, Villanova, PA 19085-1678. Villanova is a Roman Catholic university, sponsored by the Augustinian Order. An AA/EEO Employer, Villanova seeks a diverse faculty committed to scholarship, service, and especially teaching, who understand, respect, and can contribute to the University's mission and values.

324 VILLANOVA UNIVERSITY invites applications at all levels for a tenure-track position in Accounting. All research and teaching areas except Tax will be considered. The candidate should be capable of high-quality research and teaching at both the undergraduate and graduate levels. A doctorate in Accounting is required and CPA certification is preferred. Applicants for Assistant Professor positions should expect to complete doctoral requirements before the start of the contract period. Salary is competitive and commensurate with rank and carries with it the possibility of a professorship. The College of Commerce and Finance and its Accounting program are accredited by the AACSB. Representatives will be attending the AAA Annual Meeting in Hawaii. Interested candidates should send their *curriculum vitae* and an example of their scholarly work to Anthony H. Catanach, Jr., College of Commerce & Finance, Villanova University, Villanova, PA 19085-1678. Villanova is a Roman Catholic university, sponsored by the Augustinian Order. An AA/EEO employer, Villanova seeks a diverse faculty committed to scholarship, service, and especially teaching, who understand, respect, and can contribute to the University's mission and values.

325 UNIVERSITY OF VERMONT, School of Business Administration, invites applications for a tenure-track position in Accounting beginning Fall 2004. All ranks will be considered. Applicants for Assistant Professor position should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong enduring commitment to excellence in teaching and scholarly research. Teaching responsibilities include upper-level undergraduate and graduate courses in Financial and Managerial Accounting. Salary and benefits are AACSB-competitive. Members of the search committee will be attending the AAA Annual Meeting in Honolulu. Applicants should send a letter of application, current *curriculum vitae*, evidence of teaching effectiveness, and the names of three references to Professor Stephen J. Dempsey, School of Business Administration, 306 Kalkin Hall, 55 Colchester Avenue, Burlington, VT 05405. Applications will be accepted until October 15, 2003. However, we strongly encourage the submission of materials by July 31. The University of Vermont is an Affirmative Action, Equal Opportunity Employer. Applications from women and minorities are actively encouraged to apply.

326 DUQUESNE UNIVERSITY invites applications for a tenure-track position at the Assistant or Associate Professor level beginning Fall 2004. A Ph.D. or D.B.A. by May 2004 is required and professional certification is desirable. The ideal candidate is committed to teaching excellence and possesses an active research program. Teaching interest in Cost, Managerial, or Financial Accounting is preferred, but consideration will be given to all high-quality candidates regardless of expertise. Duquesne is a private, AACSB-accredited university, ranked among the top ten Catholic Universities in the United States by *U.S. News and World Report*. Duquesne's campus spans a 40-acre hilltop and is a five-minute walk from downtown Pittsburgh. There are approximately 1,000 undergraduate and 500 M.B.A. students. For consideration, submit a letter of application and current *curriculum vitae* to Professor Sharon Green, Director of Accounting Programs, A. J. Palumbo and John F. Donahue Schools of Business, Duquesne University, Pittsburgh, PA 15282. Duquesne University is an Equal Opportunity/Affirmative Action Employer.

327 SANTA CLARA UNIVERSITY, Department of Accounting, invites applications for tenure-track Assistant, Associate, or Full Professors for Fall 2004. Candidate's primary teaching interest should be Financial or Managerial Accounting. Applicants at the Assistant rank should have completed /nearly completed their Ph.D. and provide evidence of their potential for scholarly research. Candidates at higher ranks require a completed doctorate and a nationally recognized and continuing record of scholarly research and quality teaching. Review of applications will begin October 1, 2003 and continue until the positions are filled. See http://business.scu.edu/actg for more information. Candidates should send a *curriculum vitae*, two letters of recommendation, evidence of teaching effectiveness, and a sample of recent research to Suzanne Luttman, Department of Accounting, Leavey School of Business, Santa Clara University, Santa Clara, CA 95053-0380. SCU is a private Jesuit university located in Silicon Valley with competitive salaries, a housing subsidy, and summer research support available. SCU is an EEO/AA Employer.

328 NOVA SOUTHEASTERN UNIVERSITY, Wayne Huizenga School of Business and Entrepreneurship, invites applications for a position in Accounting. All ranks will be considered. Salary and benefits will be competitive. An earned doctorate in Accounting is required. A CPA or other professional certification with relevant experience is highly desirable. Successful candidates must be able to demonstrate excellent teaching skills and a scholarship commensurate with rank. Classes are held both online and on campus. NSU is among the twenty largest independent universities in the United States offering Bachelor's, Master's, and Doctoral degrees in Business, Computer Science, Dentistry, Education, Family Therapy and Studies, Law, Osteopathic Medicine, Oceanography, Optometry, Pharmacy, Psychology, and Social and Systemic Studies. The Huizenga School has approximately 2,400 working adults and pre-professional students. Applicants should include a letter describing their general qualifications, specific teaching and research interests and goals, and three references, along with a complete *curriculum vitae*, to the Chair, Search Committee, Office of the Dean, Wayne Huizenga School of Business and Entrepreneurship, Nova Southeastern University, 3100 SW 9th Avenue, Fort Lauderdale, Florida 33315. The University is an Equal Opportunity/Affirmative Action Employer.

ANNOUNCEMENT FOR TENURE-TRACK AND VISITING FACULTY POSITIONS
DEPARTMENT OF ACCOUNTING,
TIPPIE COLLEGE OF BUSINESS, THE UNIVERSITY OF IOWA

The University of Iowa Department of Accounting invites applicants for tenure-track appointments beginning Fall 2004. We will consider candidates in all areas and ranks. We will also consider one- or two-year visiting appointments if we are unable to fill the anticipated position(s) with tenure-track faculty. Required qualifications include a demonstrated interest in and capacity to do research publishable in top-tier accounting journals, a high level of teaching competence, and an earned doctorate (entry-level candidates can be near completion). We offer a competitive compensation package. We will screen applicants at all ranks and in all areas on an ongoing basis. To apply, please send a résumé, representative research output, evidence of teaching capabilities, and at least three recommendation letters to: The Recruiting Coordinator, Department of Accounting, Tippie College of Business, 108 PBB, Suite W252, The University of Iowa, Iowa City, IA 52242-1000.

The University of Iowa is an Affirmative Action, Equal Opportunity Employer. Women, minorities and individuals with disabilities are encouraged to apply.

TIPPIE HENRY B. TIPPIE COLLEGE OF BUSINESS

Nanyang Business School

Nanyang Technological University

The Nanyang Business School at Nanyang Technological University, Singapore invites applications for faculty positions in Accounting at all professional levels.

We offer undergraduate degrees in Accountancy and Business, MBA degrees, and Ph.D degrees. Our undergraduate accountancy program is highly regarded in the region, and we have a collaboration with the University of Illinois (Urbana-Champaign) in terms of our accountancy curriculum. We also co-host the International Symposium on Audit Research.

Applicants should possess a Ph.D in Accounting, and have a good record or promise in scholarly research and teaching. Salary and benefits will be competitive.

To apply, send in your application form (download from our website) or detailed CV, which should include areas of your research interest, a list of your publications, and address and e-mail of three references, to : The Vice-President (Human Resources), Administration Building, Level 4, 50 Nanyang Avenue, Singapore 639798.

E-mail : EHLEE@ ntu.edu.sg Website : http://www.ntu.edu.sg/personnel/

THE ACCOUNTING REVIEW
ISSN 0001-4826

POSTMASTER
Send Address Changes To:

AMERICAN ACCOUNTING ASSOCIATION
5717 BESSIE DRIVE
SARASOTA, FL 34233

PERIODICALS POSTAGE PAID
SARASOTA FLORIDA
AND ADDITIONAL MAILING OFFIC

34314 1 REVIEW
INTERNATIONAL BOOK HOUSE/CAL-UNIV 100
1 KYD STREET
SHOP NO 5 PALACE COURT
CALCUTTA 700 016
INDIA

P - 22

THE ACCOUNTING REVIEW

OLUME 78, No. 4 OCTOBER 2003

M BALDENIUS
Delegated Investment Decisions and Private Benefits of Control

YEESOO CHUNG and SANJAY KALLAPUR
Client Importance, Nonaudit Services, and Abnormal Accruals

OHN E. CORE, WAYNE R. GUAY, and ROBERT E. VERRECCHIA
Price versus Non-Price Performance Measures in Optimal CEO Compensation Contracts

ANDAL J. ELDER and ROBERT D. ALLEN
A Longitudinal Field Investigation of Auditor Risk Assessments and Sample Size Decisions

ARLA M. JOHNSTONE and JEAN C. BEDARD
Risk Management in Client Acceptance Decisions

ENOCK LOUIS
The Value Relevance of the Foreign Translation Adjustment

EFF L. PAYNE and WAYNE B. THOMAS
The Implications of Using Stock-Split Adjusted I/B/E/S Data in Empirical Research

RISTY L. TOWRY
Control in a Teamwork Environment—The Impact of Social Ties on the Effectiveness of Mutual Monitoring Contracts

---

QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2003–2004 Executive Committee of the Association are as follows:

President	William L. Felix, Jr., The University of Arizona
President-Elect	Jane Mutchler, Georgia State University
Vice President	Carolyn Callahan, University of Arkansas
Vice President–Education	O. Finley Graves, University of North Texas
Vice President–Finance	Judy D. Rayburn, University of Minnesota
Vice President–International	Lee Parker, University of Adelaide
Vice President–Professional Relations	Mark Chain, Deloitte & Touche
Vice President–Publications	Robert Libby, Cornell University
Vice President–Research	Zoe-Vonna Palmrose, University of Southern California
Vice President–Sections and Regions	Joanna L. Ho, University of California, Irvine
Past President	G. Peter Wilson, Boston College

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive *Accounting Education News*, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues are based on the selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.*
(This does not include Section journals.)

Select one Association-wide journal	U.S. $85.00
Select two Association-wide journals	U.S. $95.00
All three Association-wide journals	U.S. $105.00

Associate membership for students residing in the United States and Canada:

Select one Association-wide journal	U.S. $20.00
Select two Association-wide journals	U.S. $30.00
All three Association-wide journals	U.S. $40.00

Electronic journal option available for an additional U.S. $20.00 to above-mentioned full-membership and associate-membership dues.

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Artificial Intelligence/Emerging Technologies; Auditing; Financial Accounting and Reporting; Gender Issues in Accounting; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Teaching and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to taruw@u.washington.edu or sent to Professor Terry Shevlin, Senior Editor, *The Accounting Review*, School of Business, Box 353200, University of Washington, Seattle, WA 98195-3200.

The submission fee of $125.00 for members of the AAA or $200.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm and must be paid when article submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review.*

The Accounting Review (ISSN 0001-4826) is published four times a year in January, April, July, and October by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rate of U.S. $175; membership subscription rate is $25 with discounts available; and emeritus subscription rate is $8.75. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to ***The Accounting Review,*** 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2003 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Terry Shevlin
Editors
Lawrence Brown
Patricia Dechow
S. Jane Kennedy
Marlys Gascho Lipe
Madhav Rajan

VOL. 78 OCTOBER 2003 No. 4

TIM BALDENIUS
Delegated Investment Decisions and Private Benefits of Control 909
HYEESOO CHUNG and SANJAY KALLAPUR
Client Importance, Nonaudit Services, and Abnormal Accruals 931
JOHN E. CORE, WAYNE R. GUAY, and ROBERT E. VERRECCHIA
Price versus Non-Price Performance Measures in Optimal CEO Compensation Contracts 957
RANDAL J. ELDER and ROBERT D. ALLEN
A Longitudinal Field Investigation of Auditor Risk Assessments and Sample Size Decisions 983
KARLA M. JOHNSTONE and JEAN C. BEDARD
Risk Management in Client Acceptance Decisions 1003
HENOCK LOUIS
The Value Relevance of the Foreign Translation Adjustment 1027
JEFF L. PAYNE and WAYNE B. THOMAS
The Implications of Using Stock-Split Adjusted I/B/E/S Data in Empirical Research 1049
KRISTY L. TOWRY
Control in a Teamwork Environment—The Impact of Social Ties on the Effectiveness of Mutual Monitoring Contracts 1069
Editorial Data 1096
Editorial Policy and Style Information 1097
Index 1101

CENTRAL LIBRARY

Editorial Staff of *The Accounting Review*
2002–2003

SENIOR EDITOR
Terry Shevlin, University of Washington
EDITORS
Lawrence Brown, Georgia State University
Patricia Dechow, University of Michigan
S. Jane Kennedy, University of Washington
Marlys Gascho Lipe, University of Oklahoma
Madhav Rajan, Stanford University
EDITORIAL ASSISTANT
Kate Flynn Connolly

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland
COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Kathy Casper
Helen Griffith
Judy Cothern
Christine Govaars
Jan Kovarik (subcontracted)
OTHER AAA STAFF
Debora Gardner
Roland LaTulip
Deirdre Harris
Mary Cole
James Szelminski
Linda Watterworth
Barbara Brady
Myriam Chacon

Editorial Advisory and Review Board

Shannon W. Anderson, Rice University
Benjamin C. Ayers, University of Georgia
Frances L. Ayres, University of Oklahoma
William R. Baber, The George Washington University
Stephen P. Baginski, Indiana University
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Orie E. Barron, Pennsylvania State University
Eli Bartov, New York University
Anne L. Beatty, Pennsylvania State University
Robert J. Bloomfield, Cornell University
Robert M. Bowen, University of Washington
David C. Burgstahler, University of Washington
Brian J. Bushee, University of Pennsylvania
Robert M. Bushman, The University of North Carolina at Chapel Hill
C. Bryan Cloyd, University of Illinois at Urbana–Champaign
Mark L. DeFond, University of Southern California
Ilia D. Dichev, University of Michigan
Sunil Dutta, University of California, Berkeley
Merle M. Erickson, University of Chicago
Jere R. Francis, University of Missouri–Columbia
Paul Fischer, Pennsylvania State University
Jennifer Francis, Duke University
Jonathan C. Glover, Carnegie Mellon University
Wayne R. Guay, University of Pennsylvania
David A. Guenther, University of Colorado at Boulder
Mahendra Gupta, Washington University
D. Eric Hirst, The University of Texas at Austin
Christopher D. Ittner, University of Pennsylvania
James J. Jiambalvo, University of Washington
Deen Kemsley, Columbia University
William R. Kinney, Jr., The University of Texas at Austin
Lisa L. Koonce, The University of Texas at Austin
William N. Lanen, University of Michigan
Robert C. Lipe, University of Oklahoma
Joan L. Luft, Michigan State University
Russell J. Lundholm, University of Michigan
Joseph Magliolo III, Southern Methodist University
Laureen A. Maines, Indiana University
Steven R. Matsunaga, University of Oregon
Maureen F. McNichols, Stanford University
Donald V. Moser, University of Pittsburgh
James N. Myers, University of Illinois at Urbana–Champaign
Mark W. Nelson, Cornell University
Patricia C. O'Brien, University of Waterloo
Kathy R. Petroni, Michigan State University
Morton Pincus, The University of Iowa
Jamie H. Pratt, Indiana University
D. Shores, University of Washington
Richard G. Sloan, University of Michigan
Geoffrey B. Sprinkle, Indiana University
Theodore Sougiannis, University of Illinois at Urbana–Champaign
Hung-Tong Tan, Nanyang Technology University / Singapore
Richard M. Tubbs, The University of Iowa
Sandra Vera-Muñoz, University of Notre Dame
Robert E. Verrecchia, University of Pennsylvania
Beverly R. Walther, Northwestern University
Hong Xie, University of Illinois at Urbana–Champaign

PAST EDITORS

1926–1929	William A. Paton	1973–1975	Thomas F. Keller
1929–1943	Eric L. Kohler	1976–1978	Don T. DeCoster
1944–1947	A. C. Littleton	1978–1983	Stephen A. Zeff
1948–1949	Robert L. Dixon	1983–1987	Gary L. Sundem
1950–1959	Frank P. Smith	1987–1990	William R. Kinney, Jr.
1960–1962	Robert K. Mautz	1990–1994	A. Rashad Abdel-khalik
1963–1964	Lawrence L. Vance	1994–1997	Robert P. Magee
1965–1967	Wendell Trumbull	1997–2000	Gerald L. Salamon
1968–1970	Charles H. Griffin	2000–2002	Linda Smith Bamber
1971–1972	Eldon S. Hendriksen		

THE ACCOUNTING REVIEW
Vol. 78, No. 4
2003
pp. 909–930

Delegated Investment Decisions and Private Benefits of Control

Tim Baldenius
Columbia University

ABSTRACT: This paper studies the capital budgeting process in a setting where a manager is privately informed about the profitability of an investment project and enjoys nonpecuniary benefits of control ("empire benefits"). I characterize the optimal required rate of return and show that a delegation scheme with residual income-based compensation can replicate the benchmark performance achieved under centralization. The main result of the paper is that the optimal capital charge rate for computing residual income always exceeds the required rate of return as a result of empire benefits. This highlights the necessity for future empirical studies on capital budgeting to distinguish between alternative forms of hurdle rates. Contrary to conventional wisdom, I further show that if compensation contracts are derived endogenously, then the shareholders will ultimately benefit from the manager's empire benefits even under asymmetric information.

Keywords: *capital budgeting; empire-building; residual income; hurdle rates.*

I. INTRODUCTION

Capital budgeting ranks among the key managerial tasks in modern corporations and much of the attendant debate revolves around determining appropriate hurdle rates for accepting investment projects. Traditionally, the term "hurdle rate" has been used interchangeably with the required rate of return. Recently, however, a growing number of firms have adopted accounting-based performance measures known as residual income or "Economic Value Added" (EVA), which subtract a charge for the opportunity cost of capital from operating income. The capital charge rate employed is also often referred to as a hurdle rate. This dual nature of hurdle rates has remained largely unexplored in the academic literature. The present paper seeks to fill this gap by arguing that (and how) in practice the two hurdle rates should be expected to differ.

The practitioners' literature advocates using the firm's cost of capital as both the required rate of return and the capital charge rate (Brealey and Myers 2000; Young and O'Byrne 2001). On the other hand, Poterba and Summers (1995) have found that hurdle rates employed in practice significantly exceed the cost of capital. This clearly violates the

---

I thank Moshe Bareket, Zhonglan Dai, Joel Demski, Volker Laux, Nahum Melumad, Madhav Rajan, Stefan Reichelstein, Amir Ziv, two anonymous reviewers, and workshop participants at the Carnegie Mellon Accounting Conference, the FEA Conference at Rutgers University, the Duke-UNC Fall Camp, the Burton Conference, and the Finance "Free Lunch" Workshop at Columbia University, the Arijit Mukherji Conference at the University of Minnesota, and Northwestern University for helpful comments. I am especially indebted to Anil Arya and Jon Glover for discussions on this subject.

Editor's note: This paper was accepted by Madhav V. Rajan, Editor.

Submitted May 2002
Accepted May 2003

net present value rule. Such *under*investment, or capital rationing, as a result of adverse selection is consistent with shareholders curtailing informational rents earned by their privately informed managers at the expense of reduced efficiency (e.g., Antle and Eppen 1985).

At the same time, managers are frequently accused of *over*investing in assets as a consequence of private benefits they derive from controlling "empires."[1] There are numerous reasons for such empire-building tendencies to arise. Apart from intrinsic motivation, division size often is positively associated with fringe benefits such as lavish office spaces or simply power within the organization. Moreover, managers of larger and/or more profitable divisions may be more likely to be promoted or become the future CEO. Somewhat surprisingly, empire benefits have so far been largely ignored in the accounting literature.[2]

As a means to align a division manager's objective with that of the shareholders, his compensation will generally be linked to realized investment returns. While this may mitigate excessive empire-building tendencies, the manager will still invest in projects that are marginally unprofitable (and hence will lower his compensation) provided the attendant empire benefits are sufficiently high. In fact, this may be equilibrium behavior as the manager's empire benefits allow shareholders to reduce his compensation and thereby to indirectly extract a share of these benefits.[3]

This paper characterizes the capital budgeting process in the presence of adverse selection and empire benefits. In particular, I compare the optimal required rate of return with the optimal capital charge rate and benchmark both these hurdle rates against the firm's cost of capital. I refer to a firm as *centralized* if the principal makes the investment decision based on a comprehensive "revelation" contract with the better informed manager. The required rate of return then is the relevant hurdle rate under centralization. In a *decentralized* firm, the investment decision is delegated to the manager who is compensated based on residual income. The capital charge rate used for computing residual income is the relevant hurdle rate under decentralization.

Beginning with centralized firms, I obtain a familiar capital rationing result: to economize on the manager's informational rents, the firm optimally forgoes some projects that would have received funding under symmetric information. Two countervailing forces are shown to determine the comparison of the optimal required rate of return with the firm's cost of capital. While the adverse selection problem calls for a restrictive budget (i.e., a high hurdle rate), the manager's empire benefits tend to have the opposite effect: by decreasing the hurdle rate, the principal can lower the manager's compensation and thereby extract a share of the empire benefits. Depending on the relative importance of the two control problems, the optimal required rate of return can be less or greater than the cost of capital.

---

[1] Empire benefits are discussed in anecdotal form in Donaldson (1984) and Jensen (1986). *The Economist* lists empire benefits as the main example for agency problems in its online "Economics A-Z" (http://www.economist.com/research/Economics/). Burrough and Helyar (1990) describe in detail how RJR-Nabisco management wasted corporate resources by using corporate jets for their dogs, etc. Hennessy and Levy (2002) find strong empirical evidence for empire benefits to affect investment decisions. They estimate that CEOs that are members of the founding family derive an annual benefit of $3.30 per $1,000 of installed capital. Their model includes more control variables than that tested in Aggarwal and Samwick (1999), who found no evidence of empire benefits. Throughout this paper, "empire benefits" is shorthand for any investment return components that accrue directly to the manager.

[2] Exceptions are Arya et al. (1999) and Lambert (2001). Dutta (2003) considers a setting where the manager can "walk away" with the project.

[3] Baiman and Verrecchia (1995) make a related observation, albeit in a very different setting. They argue that when a firm's manager can trade on his private information, the expected gains from insider trading allow the shareholders to lower his compensation.

I then show that the principal can replicate the second-best (centralized) solution by delegating the investment decision to the manager, compensating him based on residual income. This lends support to the increasing use of residual income in decentralized firms.[4] In contrast to the ambiguous results for the required rate of return, the optimal capital charge rate under delegation always exceeds the cost of capital. This holds, because the manager as the decision maker under decentralization compares *both* return streams (cash and empire returns) with the imputed capital charge, and he ignores any informational rents, which, from the viewpoint of the principal, should be netted against the investment returns.[5] The required rate of return, in contrast, only captures the cash returns from investing. As a consequence, the main result of this paper is that the capital charge rate always exceeds the required rate of return—that is, *the hurdle rate is always higher under delegation than under centralization.*

This divergence in the hurdle rates is in stark contrast to the practitioners' literature cited above. The findings in this paper indicate that lumping together required rates of return with capital charge rates, as is common in empirical studies on capital budgeting practices, adds noise. This may help explain the somewhat inconclusive cross-sectional findings on hurdle rates in Poterba and Summers (1995). Based on my results, a required rate of return of, say, 15 percent will result in lower levels of investment than an equally high capital charge rate in a decentralized, but otherwise identical, division or firm. This is further corroborated by additional results that predict qualitatively different comparative statics for the two hurdle rates.

The business press generally views empire-building tendencies as value-destructing because of the resulting investment distortions. As the preceding discussion suggests, this need not be true once compensation contracts are determined endogenously. Indeed, an immediate corollary of my results is that the firm owners can always extract a sizable share of the manager's empire benefits and, hence, will ultimately benefit from an exogenous increase in these benefits. It should be noted, however, that this result hinges on the assumption that empire benefits depend on the characteristics of the investment project in a deterministic and commonly known fashion. Put differently, for the negative connotation associated with empire benefits to be justified, it must therefore be the case that shareholders are imperfectly informed about how these benefits are driven by the scale and profitability of the assets under a manager's control.[6]

This paper is related to several recent studies on capital budgeting. Christensen et al. (2002) and Dutta and Reichelstein (2002) examine optimal hurdle rates in the absence of empire benefits.[7] Closest to my model is Lambert (2001), who studies a capital budgeting model with asymmetric information and empire benefits. While Lambert (2001) does not allow for screening of the manager's private information, this paper derives suitable hurdle rates that implement the optimal revelation mechanism. All three earlier studies emphasize the role of project (or market) risk in that any deviations between required rates of return and capital charge rates are solely driven by differences in risk tolerance between shareholders and managers.[8] While corporate finance theory predicts that the required rate of return is, in general, increasing in a project's riskiness, Christensen et al. (2002) demonstrate that this need not hold for the capital charge rate under delegation. A risk-averse manager will internalize the project risk via the attendant increase in his compensation risk for a given capital charge rate. Raising the hurdle rate further would thus amount to "double-counting" of the risk premium.[9] In fact, Poterba and Summers (1995, 47) find that none of the traditional risk factors correlates with hurdle rates in their sample. By introducing empire benefits, the present paper takes the natural next step in this research agenda as it demonstrates systematic differences between the two hurdle rates even in a risk-neutral setting.[10]

---

[4] Note that in my model the compensation contract under delegation does not require any communication between manager and principal, while this is necessary for the centralized solution. Since I do not explicitly model any communication costs, however, the revelation principle implies that a decentralized firm can at best replicate the performance of a centralized firm.

[5] As shown below, under delegation the manager internalizes 100 percent of the empire returns, but only a fraction $\beta < 1$ of the cash returns (and outlay) from investing. Here, β is the bonus coefficient that is endogenously determined by the severity of the agency problem (i.e., by the marginal disutility of effort). The more severe the agency problem, the higher the bonus coefficient and, therefore, the less weight the manager assigns to the empire benefits relative to his utility from monetary compensation.

[6] Formally addressing this issue would require a two-dimensional adverse selection model where the manager privately observes k and a measure of his empire benefits, where the latter may or may not be correlated with k. However, the theory of such multidimensional screening problems is highly incomplete at this point; see Rochet and Stole (2001) for a survey of the current state of this literature.

The remainder of the paper is organized as follows. Section II describes the model. Section III characterizes the optimal centralized capital budgeting procedure and the corresponding required rate of return. Section IV addresses delegated investment decisions and capital charge rates. Section V analyzes the comparative statics. Section VI concludes.

II. MODEL

This section develops a single-period capital budgeting model of a firm consisting of a risk-neutral principal and a risk-neutral manager. The model entails private precontractual information (adverse selection), unobservable effort choices, and private benefits to the manager from controlling assets. At date 0, the firm can invest in a project, with $I \in \{0, 1\}$ as an indicator variable that denotes whether the project is undertaken. If it is ($I = 1$), then the attendant expense equals b. At date 1, the return from the investment is realized. The gross cash return from investing equals $k \in [\underline{k}, \bar{k}]$ and is known only to the manager prior to accepting the contract. The manager exerts personally costly effort, denoted by $a \in [0, \bar{a}]$. The (commonly known) marginal productivity of effort is normalized to 1, without loss of generality. Hence the date–1 cash flow equals $c = kI + a$. The principal maximizes the total date–0 value of the firm net of the manager's compensation, s:

---

[7] Baldenius and Ziv (2003) incorporate corporate income taxes into a capital budgeting setting. Their model, which builds on Rogerson (1997) and Reichelstein (1997), predicts that firms employing pre-tax residual income tend to set higher capital charge rates than comparable firms using after-tax metrics such as "Economic Value Added."

[8] Lambert (1986) analyzes the investment distortions that arise if a risk-averse manager has to be motivated to exert personally costly effort to become informed about a project's profitability.

[9] Christensen et al. (2002) distinguish between firm-specific and market risk. They argue against using the weighted-average cost of capital (WACC) as the capital charge rate and show that only firm-specific (diversifiable) risk affects the capital charge rate, whereas market risk does not. As a result of firm-specific risk, the hurdle rate can be less or greater than the risk-free rate depending on the information held by the manager at the time of investing.

[10] Bernardo et al. (2001) also study the role of compensation to control managers' empire-building tendencies, albeit only in a centralized decision-making setting. Their analysis is thus confined to required rates of return as hurdle rates, for which they derive results consistent with the earlier capital rationing literature, e.g., Harris et al. (1982), and Antle and Eppen (1985). In Harris and Raviv (1996, 1998), headquarters uses internal auditing procedures, instead of compensation, to mitigate its divisional managers' desire for empires.

$$\gamma[kI + a - s] - bI.$$

Here, $\gamma \equiv (1 + r)^{-1}$ is the firm's discount factor and r is the cost of the capital, which in this risk-neutral setting equals the risk-free rate. The net present value of the investment alone hence is $NPV(k|r) = \gamma k - b$.

In a key departure from traditional adverse selection models, I assume that the manager enjoys *empire benefits*—or private benefits of control—from overseeing the project. As a consequence, the investment project has two return components: one accruing directly to the principal (the cash return) and one accruing directly to the manager (the empire return). The empire benefits, denoted by $e(k,I)$, are nondecreasing in the profitability of the project and in the amount invested. The manager's utility, evaluated at date 1, comprises his compensation and empire benefits less his disutility of effort, $v(a)$:

$$U = s + e(k,I) - v(a),$$

with $v'(a) > 0$, and $v''(a) \geq 0$. For ease of exposition (none of the results will depend on this), $e(\cdot)$ is assumed to be multiplicatively separable in that $e(k,I) = e(k)I$, where $e'(k) \geq 0$ and $e''(k) \leq 0$. When the contract is signed, the principal, in contrast to the manager, knows only the distribution function $F(k)$, with a corresponding density $f(k)$, from which k is drawn. As is common in adverse selection models, the inverse hazard rate $H(k) \equiv [1 - F(k)]/f(k)$ is assumed to be decreasing.

As a benchmark, I begin by characterizing the *first-best* investment decision that would result if the principal could observe k and thereby could also infer the manager's effort choice a.[11] The effort and investment problems are then effectively separated. The principal reimburses the manager for his disutility of effort and invests according to: $\hat{I}(k) \in \arg\max_I\{\gamma[kI + e(k)I] - bI\}$. The project would then be accepted if, and only if, $k \geq \hat{k}$, where:

$$\gamma[\hat{k} + e(\hat{k})] - b = 0. \tag{1}$$

Hence, under symmetric information the investment would be undertaken whenever the sum of the two return components (cash and empire benefits), in present value terms, exceeds the initial investment—that is, whenever $NPV(k|r) + \gamma e(k) \geq 0$.

The practitioners' literature cited in the "Introduction" offers various explanations for empire benefits to arise, apart from intrinsic motivation. First, investments often entail hiring additional staff. Divisional managers may thereby gain more power or simply enjoy more lavish office spaces and other fringe benefits. Second, the manager's future career opportunities may improve if the projects under his control turn out to be profitable.[12] While in the first case, empire benefits will be largely independent of the project's profitability, this is not plausible in the second case. Here, empire benefits might well be negligible for low values of k, but increasing in k. To address these differences more formally, I shall invoke

---

[11] Since both the manager and the principal are risk-neutral, all results for the adverse selection model remain valid if the date–1 cash flow were subject to an additional random noise term: $\tilde{c} = kI + a + \tilde{\varepsilon}$ with $E[\tilde{\varepsilon}] = 0$. For the first-best benchmark setting this would require the principal to be able to directly observe a, as the latter could otherwise no longer be inferred from c.

[12] Note that the present model does not explicitly incorporate such dynamic features; see instead Holmstrom and Ricart i Costa (1986).

$$U(\tilde{k},k) = s(\tilde{k}) + e(k)I(\tilde{k}) - v(a(\tilde{k},k)). \tag{3}$$

Let $U(k) \equiv U(k,k)$. The revelation principle (e.g., Myerson 1979) applies and, hence, the search for optimal contracts can be confined to those that induce truthful reporting of k. Normalizing the manager's date–1 market alternative to zero, without loss of generality, the principal's problem becomes:[16]

$$\begin{aligned} \mathbf{P}_0\text{:}\ & \max_{I(\cdot),a(\cdot),s(\cdot)} \ E_k\{\gamma[kI(k) + a(k) - s(k)] - bI(k)\}, \\ & \text{subject to: } U(k,k) \geq U(\tilde{k},k), \text{ for all } k,\tilde{k}, \\ & \qquad\qquad\quad U(k,k) \geq 0, \text{ for all } k. \end{aligned} \tag{4}$$

According to Program $\mathbf{P}_0$, the principal maximizes the date–0 value of the firm, subject to an incentive compatibility constraint ensuring the manager reports truthfully and a participation constraint providing the manager with his reservation utility.

In contrast to the vast body of the adverse selection literature, the manager's private information parameter, k, enters his utility function in (3) twice, resulting in countervailing reporting incentives. Underreporting $\tilde{k} < k$ results in a less demanding cash flow target, which the manager can achieve with less effort. On the other hand, overreporting $\tilde{k} > k$ might trigger the investment decision, resulting in empire benefits. Unlike earlier studies on countervailing incentives (e.g., Lewis and Sappington 1989; Maggi and Rodriguez-Clare 1995), however, here the manager unambiguously benefits from a higher realization of the *true* profitability, k, because $e'(k) \geq 0$ and $a(\tilde{k},k)$ is decreasing in k. That is, the manager consumes higher empire benefits and he can achieve a given cash flow target with less effort. As a consequence, standard techniques for solving adverse selection models suffice.[17]

For any feasible solution to $\mathbf{P}_0$ satisfying individual rationality and "local" incentive compatibility, i.e., $\partial U/\partial \tilde{k} = 0$, the manager's informational rent equals (see proof of Lemma 1 in Appendix A for details):

$$U(k) = U(\underline{k}) + \int_{\underline{k}}^{k} [e'(u) + v'(a(u,u))]I(u)du. \tag{5}$$

Provided, as assumed above, it is always worthwhile to induce maximum managerial effort, by using (5) the isolated investment problem becomes:[18]

$$\begin{aligned} \mathbf{P}_1\text{:}\ & \max_{I(k)} E_k\{(\gamma[k + e(k) - (e'(k) + v')H(k)] - b)I(k)\}, \\ & \text{subject to: } I(k) \text{ nondecreasing in } k, \end{aligned} \tag{6}$$

where $v' \equiv v'(\overline{a})$, to simplify notation. The principal's problem thus amounts to maximizing the expected date–0 gross firm value cum empire benefits, less the manager's informational rents.

---

[16] As pointed out by an anonymous reviewer, the empire benefits turn this adverse selection problem into one with type-dependent reservation utilities of $\overline{U} - e(k)I$ where $\overline{U} = 0$; see also Jullien (2000).

[17] Note that this may not be the case in a model with continuous investments where, for some realizations of k, the over- and underreporting incentives may just offset each other.

[18] The monotonicity constraint in Program $\mathbf{P}_1$ ensures that the solution found by observing only the "local" incentive constraints also satisfies "global" incentive compatibility, as required in Program $\mathbf{P}_0$.

The solution to Program $\mathbf{P}_1$ calls for the firm to invest in the project if, and only if, $k \geq k^*$, where the optimal profitability cutoff k^* is given by:

$$\gamma[k^* + e(k^*) - (e'(k^*) + v')H(k^*)] - b = 0. \tag{7}$$

My first result is a variant of the standard capital rationing finding in the adverse selection literature (e.g., Antle and Eppen 1985).

Lemma 1: The second-best centralized contract solves Program $\mathbf{P}_1$ and is characterized by underinvestment: for $k \in (\hat{k}, k^*)$ the firm forgoes projects that would be undertaken if k were publicly observed.

Proof: All proofs are contained in Appendix A.

Capital rationing arises endogenously as a consequence of the principal's trade-off between maximizing the gross firm value cum empire benefits and minimizing the manager's informational rents. For any $k < k^*$, the manager is compensated for his disutility of effort and just meets his reservation utility. If k exceeds the cutoff and the project is undertaken, then (as shown in Appendix A) the manager earns informational rents of:

$$U(k > k^*) = (k - k^*)v' + e(k) - e(k^*). \tag{8}$$

Note that Lemma 1 does not refer to the forgone projects as "profitable" in the classical corporate finance sense. They may well have a negative net present value as long as they are associated with offsetting private benefits of control. In the optimal centralized solution, the manager's compensation comprises his disutility of effort plus his informational rent, net of his empire benefits. Specifically, at the second-best cutoff, $s(k^*) = v(\bar{a}) - e(k^*)$.[19]

Instead of determining cutoffs for absolute investment returns, k, firms in practice tend to set *required rates of return*. Empirical studies have shown that these relative profitability hurdle rates consistently exceed the firms' cost of capital (e.g., Poterba and Summers 1995). As noted above, this is consistent with shareholders curtailing managers' informational rents. The natural next step now is to ask whether this remains true once private benefits of control are accounted for.

Let the internal rate of return of an investment project with profitability k be denoted by $R(k)$. Thus, $NPV(k|R(k)) \equiv 0$ or, equivalently, $R(k) \equiv k/b - 1$. Consistent with the traditional textbook notion of internal rates of return (Brealey and Myers 2000), $R(k)$ captures *only the cash returns* from investing, ignoring any nonpecuniary empire benefits. Note that there is a one-to-one correspondence between k and $R(k)$: instead of having the manager report k, an equivalent revelation mechanism exists where he is asked to report $R(k)$. From (1), it follows that the first-best required rate of return in the absence of informational asymmetries is less than the firm's cost of capital, i.e., $R(\hat{k}) \leq r$, since the principal can fully extract the manager's empire benefits. The second-best required rate of return that corresponds to k^*—as given in (7)—equals:

$$R^* \equiv R(k^*) = \frac{k^*}{b} - 1. \tag{9}$$

---

[19] Abstracting from empire benefits, the difference between the manager's compensation and his disutility of effort corresponds to the "organizational slack" in Antle and Eppen (1985).

Recall that there are two control problems present: the manager enjoys private benefits of control, and he can reduce his effort input by underreporting the investment return (the agency problem). These two control problems have countervailing effects on the optimal required rate of return. Holding constant the manager's empire benefits, the importance of the agency problem is determined by ν' and by the *ex ante* profitability of the project as parameterized by b. The lower b, the more profitable the project is in the expectation (all else equal) and, hence, the higher the probability that the investment is undertaken and that the manager earns informational rents.

The first main result compares the second-best required rate of return with the cost of capital for the plush office and reputation scenarios introduced in Definition 1.

Proposition 1: Suppose the principal makes the investment decision based on a required rate of return, R^*.

1. If (PO) holds, then for any (b, ν') there exists a unique cutoff $\hat{e}_0 \geq 0$, such that $R^* < r$ *if* $e_0 > \hat{e}_0$, while $R^* > r$ *if* $e_0 < \hat{e}_0$.
2. If (REP) holds, then for any (ν', e_1) there exists an initial investment expense, $\hat{b}$, such that $R^* > r$ holds for all $b < \hat{b}$.

Empire benefits always result in a relaxation of the manager's participation constraint. In the plush office (PO) scenario, these benefits are independent of k and hence they cannot be used to screen different k-types. As a consequence, for high e_0 the principal invests in negative-NPV projects while reducing the manager's compensation.[20] On the other hand, as e_0 becomes small, the model converges to the standard adverse selection model and I recoup the classical capital rationing result that $R^* > r$.

In the reputation (REP) scenario, the empire benefits are informative about k. Thus, they can be used to screen the manager's information more effectively. Technically speaking, empire benefits now also affect the manager's incentive constraint. If the project is very profitable in expectation (b is low) and the agency problem is severe (ν' is high), then rent extraction weighs heavily in the principal's objective function for two reasons: the manager earns rents for many realizations of k, and these rents are increasing "steeply" in k. Therefore, the optimal capital budgeting policy tends to be more restrictive in that $R^* > r$ holds for any e_1.[21] This is in contrast to the plush office scenario (Proposition 1, part 1) where $R^* < r$ always holds for sufficiently high empire benefits.

To make these trade-offs more transparent, I consider a specific example where k is uniformly distributed. This allows for closed-form solutions to be obtained.

Example: Suppose $k \sim U[0, \bar{k}]$ so that $H(k) = \bar{k} - k$. Rewriting (7) for the case where $e(k) = e_0 + e_1 k$, as in (PO) and (REP), then gives:

$$k^* = \frac{(1 + r)b - e_0 + (e_1 + \nu')\bar{k}}{1 + 2e_1 + \nu'}. \tag{10}$$

---

[20] "High" empire benefits, e_0, here means high relative to the agency problem parameter, ν'. Note that $\nu' \leq 1$ implicitly holds by the assumption that maximum effort is always desirable.

[21] Note that no explicit lower bound has been imposed on the initial investment expense, b. In fact, b can even be ("slightly") less than $\gamma \underline{k}$ without making the investment problem under adverse selection trivial. Even then the principal may optimally forgo a project with a k-realization close to $\underline{k}$ so as to extract informational rents from the manager.

The expression in (10) can in turn be rewritten to yield $r = [k^* + \delta(k^*)]/b - 1$, for $\delta(k) \equiv e_0 + (2k - \bar{k})e_1 - (\bar{k} - k)v'$. By comparison with (9), $R^* \equiv k^*/b - 1 > r$, if $\delta(k^*) < 0$, and vice versa.[22]

Consider the plush office (PO) scenario where $e_1 \rightarrow 0$. Using (10), I find that $\delta(k^*) < 0$ (and hence $R^* > r$) if, and only if, $e_0 < \hat{e}_0 = [\bar{k} - (1 + r)b]v'$. Note that $\hat{e}_0 > 0$ always holds. Thus, R^* can indeed be less or greater than r for any (v',b), depending on e_0. Now consider the reputation (REP) scenario where $e_0 \rightarrow 0$. Then, for b sufficiently small in that $(1 + r)b < \frac{1}{2}\bar{k}$, $\delta(k^*) < 0$ holds if, and only if, $e_1 < v'[\bar{k} - (1 + r)b]/[2(1 + r)b - \bar{k}] < 0$. This contradicts the earlier assumption that $e_1 \geq 0$. Hence, if $b < \hat{b} = \frac{1}{2}\gamma\bar{k}$ in the (REP) scenario with uniformly distributed types, then $R^* > r$ holds irrespective of the magnitude of empire benefits.

To summarize, the optimal required rate of return in connection with centralized investment decisions can be greater or less than the cost of capital, depending on the magnitude and the functional form of the manager's empire benefits. An unambiguous ranking is obtained only if empire benefits are driven by project scale and profitability—that is, in the reputation (REP) setting—and if the project is relatively profitable, *ex ante*. In this case, the hurdle rate always exceeds r.

IV. DELEGATED INVESTMENT DECISIONS

The capital budgeting model analyzed so far was characterized by a high degree of centralization. In practice, however, capital budgeting procedures tend to be more decentralized. I therefore now consider a regime where the manager makes the investment decision and is compensated based on residual income. First, residual income (or EVA) has been adopted by a large number of firms and, second, recent studies have shown that this metric has desirable features for delegating investment decisions (e.g., Rogerson 1997; Reichelstein 1997).

I consider linear contracts based on residual income (RI), which equals divisional income less a charge for the capital invested. In this one-period setting:

$$RI(k,a,\rho) = kI + a - (1 + \rho)bI, \tag{11}$$

with ρ denoting the *capital charge rate*. The compensation contract is given by:

$$s(RI) = \alpha + \beta \times RI,$$

where α denotes fixed salary and β is the bonus coefficient. Residual income is said to be *optimal*, if such a linear compensation contract implements the second-best (centralized) effort and investment decisions and results in the second-best expected firm profit net of compensation.[23] It now becomes apparent that two alternative interpretations of "hurdle rates" are conceivable: as the required rate of return $R(k) = k/b - 1$, or as the capital charge rate, ρ, under delegation.

---

[22] Consistent with results in Section V, the following comparative statics hold for the optimal cutoff in (10): k^* is decreasing in e_0, increasing in v' and b, while the sign of $\partial k^*/\partial e_1$ is ambiguous.

[23] Note that the linear compensation schemes considered here do not require any communication between the principal and the manager. This would no longer be true if the optimal solution were to entail interior effort levels (see Appendix B).

A manager compensated based on residual income will choose:

$$\max_{a,I} \{\alpha + \beta \times RI(k,a,\rho) - v(a) + e(k)I\}. \tag{12}$$

Suppose the principal sets the bonus coefficient equal to the manager's marginal disutility of effort at the desired effort level, i.e., $\beta = v'$. Then the manager will invest whenever:

$$v'[k - (1 + \rho)b] + e(k) \geq 0. \tag{13}$$

The next result shows that a delegation scheme can replicate the benchmark performance of centralization and it characterizes the appropriate hurdle rate.

Proposition 2: Residual income constitutes an optimal performance measure. The optimal capital charge rate exceeds the firm's cost of capital; that is, $\rho^* > r$.

The proof of Proposition 2 proceeds by constructing contract parameters:

$$\beta = v' \qquad \text{and} \qquad \alpha = v(\bar{a}) - v'\bar{a}, \tag{14}$$

and an appropriate capital charge rate, for which the manager will take the desired actions and earn the second-best informational rent. As shown in Appendix A, the manager's informational rent under such a delegation scheme equals that under centralization, which is given in (8) above.

To see why the capital charge rate, ρ^*, which together with the contract parameters in (14) achieves optimality, has to exceed the cost of capital, divide the condition that characterizes the second-best cutoff k^*—Equation (7)—by the discount factor γ. This yields (15) below. Then compare this with the condition that describes that manager's investment decision under delegation—Equation (13)—evaluated as an equality at $k = k^*$ and divided by the bonus coefficient $\beta = v'$. This gives (16):

$$k^* + e(k^*) - (e'(k^*) + v')H(k^*) - (1 + r)b = 0, \tag{15}$$

$$k^* + \frac{1}{v'} e(k^*) - (1 + \rho^*)b = 0. \tag{16}$$

At the margin, the manager trades off cash returns versus empire returns: he is willing to accept a decrease of \$1 in the performance measure for an additional \$$(v')$ of nonpecuniary empire benefits. Comparing (15) with (16), it is apparent that $\rho^* > r$ holds for two reasons: (i) the empire benefits are scaled by the bonus coefficient, and (ii) the manager does not subtract any informational rents when choosing the cutoff.

While Proposition 2 only lays out *an* optimal delegation scheme, this solution can indeed be shown to be unique within the class of residual income-type performance metrics. Note that the second-best investment and effort choices can also be implemented by a different delegation scheme with $\beta \in [v',1]$, $\alpha = v(\bar{a}) - \beta\bar{a}$ and a hurdle rate, $\rho(\beta)$, chosen as a plug variable so that $k^* - [1 + \rho(\beta)]b + e(k^*)/\beta = 0$ holds. For this equality to hold for any $\beta \in [v',1]$, it must be that $\rho'(\beta) < 0$, yet $\rho(1) > r$. However, the manager's informational rent for any $k > k^*$ then equals $(k - k^*)\beta + e(k) - e(k^*)$, which exceeds that in (8) under centralization. Thus, setting $\beta > v'$ requires paying the manager excessive informational rents while setting $\beta < v'$ results in suboptimal effort input. Therefore, $\beta = v'$ characterizes the *unique* optimal solution. It is instructive to contrast these findings with Lambert (2001), who does not consider information-eliciting mechanisms: since the principal does not care about rent extraction in his model, the capital charge rate equals r and $\beta = 1$ holds for a risk-neutral manager.

It is also important to note that the prediction of slope coefficients $\beta < 1$ in my model is not an artifact of the assumed corner solution for effort choices. As shown in Appendix B, $\beta(k) < 1$ will hold for all k greater than the investment cutoff, even in the case of interior optimal effort choices (and thus type-dependent slope coefficients). A more steeply sloped contract (i.e., a higher value of β) always implies higher informational rents for the manager. At the margin, the principal therefore is willing to accept suboptimal effort levels so as to economize on the manager's rents.

A comparison of Propositions 1 and 2 shows that the optimal hurdle rate under centralization, R^*, can be either greater or less than the cost of capital, whereas the capital charge rate under delegation, ρ^*, always exceeds r. This holds *because the manager as the decision maker under delegation takes into account both cash and empire returns from investing*. In contrast, the required rate of return, by definition, ignores the empire returns. Again rewriting (16) as:

$$\rho^* = \frac{k^*}{b} + \frac{1}{v'}\frac{e(k^*)}{b} - 1, \tag{17}$$

yields the following result by a direct comparison with the second-best required rate of return in (9). It is therefore stated without formal proof.[24]

Proposition 3: The hurdle rate under delegation always strictly exceeds that under centralization—that is, $\rho^* > R^*$.

The results thus far demonstrate that the empirical findings reported in Poterba and Summers (1995) have to be viewed in the broader context of organizational design—in particular, the contracting environment. The above arguments imply that, for a given investment opportunity set, a required rate of return of, say, 15 percent would lead to a lower level of investment than a 15 percent capital charge rate under residual income-based compensation. Put differently, the two hurdle rates will differ in their economic consequences as a result of private benefits of control.

V. COMPARATIVE STATICS

This section addresses in more detail how empire benefits affect both kinds of hurdle rates. This discussion sheds light on two related questions. First, how does the optimal solution characterized above compare with the optimal capital budgeting mechanism in the absence of empire benefits? Second, do empire benefits invariably harm shareholders, as suggested by conventional wisdom?

To address these questions, I introduce a new and commonly known parameter, $m \in [0,\bar{m}]$, which captures the *scale* of the manager's empire benefits so that:

$$e(k,I,m) \equiv me(k)I. \tag{18}$$

---

[24] To avoid the trivial case where $e(k^*) = 0$ and hence $\rho^* = R^*$, I assume that $e(k) > 0$ for small k.

For instance, Hennessy and Levy (2002) measure the propensity of managers to engage in empire-building by: (i) whether the manager is a member of the founding family; (ii) the percentage of insiders on the board of directors; and (iii) the manager's tenure. In their sample, managers who are members of the founding family are shown to consume higher empire benefits, while (ii) and (iii) are used as proxies for the likelihood that the manager's decision will be challenged by the board. Also, for the reputation setting (REP), one would expect m to be relatively high when the manager enjoys high external visibility (e.g., due to high analysts' following). The preceding analysis therefore has captured the special case of $m = 1$.

Upon substituting $e(k, I, m)$ for $e(k, I)$, let $k^*(m)$ denote the solution to the centralized problem as in (7) for any m, and $R^*(m) \equiv R(k^*(m))$. Similarly, let $\rho^*(m)$ be the optimal capital charge rate under delegation when generalizing the solution characterized in (17). A first intuition suggests that $k^*(m)$ be decreasing in m as a consequence of the manager's compensation/empire benefits trade-off. While this intuition holds for the symmetric information case, it may fail under adverse selection, as the next result shows.[25]

Proposition 4: Suppose the manager's empire benefits are as given in (18).

1. If (PO) holds, then $R^*(m)$ is decreasing in m.
2. If (REP) holds, then $R^*(m)$ is increasing in m, for b sufficiently small.
3. In a decentralized firm with residual income-based compensation, the optimal capital charge rate, $\rho^*(m)$, is always increasing in m.

In the plush office scenario (PO) the intuition developed for the symmetric information case applies: an increase in m has the sole effect of relaxing the manager's participation constraint. The principal then lowers the required rate of return R^* and at the same time reduces the manager's compensation. In the reputation setting (REP), an increase in m again also allows for improved screening of different k-types. If the project is sufficiently profitable in expectation (b is small), then this incentive constraint-effect dominates and the firm raises R^* so as to reduce the rents earned by high-k managers.[26]

Under delegation, two effects are present. First, as m increases, the optimal profitability cutoff, $k^*(m)$, to be implemented via the capital charge rate, changes. As the preceding paragraph has shown, this effect can go either way. Second, the manager becomes more eager to invest irrespective of the project's profitability. As the proof demonstrates, this second, more direct, effect always dominates the first one so that $\rho^*(m)$ is always increasing. Note that the contract parameters (α,β) given in (14) are independent of the manager's empire benefits. Instead, the hurdle rate $\rho^*(m)$ is the sole instrument the principal uses to extract (a share of) these benefits.

Private benefits of control are commonly perceived as a "public bad" in the business press. The claim is that managers' inclination to overinvest relative to the NPV rule destroys firm value. As the previous analysis suggests, this need not be the case once compensation contracts are derived endogenously. In fact, the delegated investment condition in (13) reveals that shareholders in this model will ultimately benefit from their manager's empire benefits even in the presence of adverse selection. For any $m > 0$, the shareholders can always induce the manager to invest whenever $k \geq k^*(0)$ by simply raising ρ. Holding the contract parameters (α,β) fixed, this will transfer rents from the manager to the shareholders while ensuring all incentive and participation constraints remain intact. By revealed preference, the shareholders can do even better by choosing $\rho^*(m)$ optimally. This discussion is summarized in the following corollary.

---

[25] Note that, while parts 1 and 2 of Proposition 4 rely on linear specifications for the $e(k)$-function, this is not the case for part 3.

[26] Part 2 of Proposition 4 complements a comparative statics result derived in Bernardo et al. (2001). In their model, investments are continuous and effort and investment are complements. Their Proposition 4 part (i), shows that (pointwise) investment amounts are increasing in empire benefits for high types and decreasing for low types.

Corollary 1: Suppose the manager's empire benefits are as given in (18). Then the value of the firm, net of the manager's compensation, is increasing in m.

Rather than concluding that the conventional wisdom of empire benefits destroying firm value is false, it should be noted that Corollary 1 hinges on the assumption that a project's financial and nonpecuniary returns are related in a deterministic (and commonly known) fashion. In practice, this may not always be the case. Put differently, the corollary shows that a necessary condition for empire benefits to reduce firm value is that shareholders have only incomplete information about how these benefits depend on the size and profitability of investment projects (see also footnote 6).[27]

VI. CONCLUSION

This paper has studied optimal hurdle rates in a setting where investment decisions are subject to adverse selection and private benefits of control. The results have demonstrated systematic differences between the required rate of return and the capital charge rate, which are the two most common incarnations of hurdle rates. A key empirical implication of the model is that capital charge rates used for residual income/EVA-based bonus plans should on average be higher than required rates of return across comparable firms. Also, capital charge rates should be higher in high-empire-benefits environments (e.g., a manager has been with the firm for a long time; high percentage of insiders among the board members; high external visibility) as compared with low-empire-benefits environments. This need not be true for required rates of return.

The risk-neutral framework adopted in this paper was in part motivated by Poterba and Summers' (1995) finding of hurdle rates unrelated to measures of firm-specific risk. In light of Christensen et al. (2002) and Dutta and Reichelstein (2002), it would be desirable to reassess firms' choice of hurdle rates in a large-sample study, taking into account firm-specific and market-wide risk factors, as well as proxies for empire benefits.

The results in this paper have demonstrated that different assumptions regarding the specific nature of empire benefits have qualitatively different implications for the capital budgeting process. This has been ignored in previous studies. As shown here, the optimal hurdle rates depend on whether empire benefits are driven by project size and profitability, or by size only. From an empirical standpoint, this raises the question of how to distinguish between these two scenarios. One possible control variable that comes to mind is a manager's tenure. Near retirement, managers tend to put less emphasis on future career prospects and hence may prefer larger projects over smaller, but more profitable ones.[28] The opposite is likely to hold at an early stage in a manager's career where reputation effects play a more important role.

---

[27] In a similar vein, Corollary 1 is sensitive to the specifics of the manager's utility function. In my model, the manager views compensation and empire benefits as substitutes. Alternative specifications are conceivable—e.g., lexicographic preferences—for which this result may not hold.

[28] Gibbons and Murphy (1992) analyze investment levels as a function of managers' tenure and find investments in R&D to increase toward the end of managers' careers, which is inconsistent with maximizing earnings-based compensation. See also Hennessy and Levy (2002). However, neither of these studies tests for *ex post* profitability of investment projects.

The present model can be generalized to analyze capital assets with a useful life of multiple periods, in which case accrual accounting issues arise. In the absence of empire benefits, this question of how to depreciate an asset so as to achieve intertemporal matching of investment costs with investment returns has been the subject of several recent studies.[29] These studies have argued that depreciation for projects with uniform cash flows over time should be decelerated—that is, increasing over time. This is clearly at odds with firm practice, as companies predominantly use straight-line (or even faster) depreciation. In the presence of empire benefits, intertemporal matching has to take both cash and nonpecuniary return streams into considerations.[30] In particular, it is plausible that empire benefits are particularly high in early periods—e.g., due to analysts' coverage. If this is the case, then, as shown in an earlier version of this study (Baldenius 2002), optimal depreciation tends to be more accelerated than previously suggested. Hence, empire benefits may help explain depreciation schedules observed in practice, and the recent call for more decelerated depreciation to overcome perceived underinvestment problems (Young and O'Byrne 2001) may thus be misguided.

APPENDIX A
PROOFS

Proof of Lemma 1

I begin by stating a sufficient condition for the principal to induce the manager to exert maximum effort for any realization of k, i.e., for $a(k) \equiv \bar{a}$:

$$1 - v'(\bar{a}) - v''(\bar{a})H(\hat{k}) \geq 0, \tag{19}$$

where $H(k) \equiv [1 - F(k)]/f(k)$. Here, $\hat{k}$ denotes the first-best profitability cutoff, i.e., the profitability parameter for which the project just breaks even, i.e., $\gamma[\hat{k} + e(\hat{k})] = b$. I will show below that (19) is indeed sufficient for $a(k) \equiv \bar{a}$, for all k, in equilibrium.

The manager's date–1 utility if he reports $\tilde{k}$ while actually being of type k equals $U(\tilde{k},k) = s(\tilde{k}) + e(k)I(\tilde{k}) - v(a(\tilde{k},k))$. An increase in the manager's true type, k, has a positive impact on his utility (i) due to higher empire benefits and (ii) because it will allow him to achieve a given cash flow objective with less effort, i.e., $a(\tilde{k},k)$ is decreasing in k since, by (2), $a(\tilde{k},k) = c(\tilde{k}) - kI(\tilde{k})$. Hence, for any contract that satisfies the participation and (local) incentive compatibility constraints, the envelope theorem implies that $U'(k) = [e'(k) + v'(a(k,k))]I(k) \geq 0$. As a consequence, the manager's individual rationality constraint will be binding for $k = \underline{k}$ and his informational rent for any k equals:

$$U(k) = U(\underline{k}) + \int_{\underline{k}}^{k} [e'(u) + v'(a(u,u))]I(u)du.$$

In the optimal solution, the lowest type, $\underline{k}$, will earn zero rents. In expected terms, this implies $E_k\{U(k)\} = \int_{\underline{k}}^{\bar{k}} [e'(k) + v'(a(k,k))]I(k)H(k)f(k)dk$, after integration by parts. Hence:

---

[29] See Rogerson (1997), Reichelstein (1997), Bareket (2002), and Dutta and Reichelstein (2002).

[30] When addressing type-dependent empire benefits in form of a multiperiod "reputation" scenario, it is conceivable that the manager's reservation utility in future periods is endogenously determined by the realization of k. The principal would then have to take this into account when designing the optimal capital budgeting process, i.e., the optimal investment cutoff.

$$E_k\{s(k)\} = \int_{\underline{k}}^{\bar{k}} [\nu(a(k,k)) - e(k)I(k) + [e'(k) + \nu'(a(k,k))]I(k)H(k)]f(k)dk.$$

Recall from (2) that $a(k,k) = c(k) - kI(k)$. Substituting the last two expressions into the principal's objective in $\mathbf{P}_0$ yields:

$$\max_{I(k),c(k)} \int_{\underline{k}}^{\bar{k}} \{\gamma[c(k) + e(k)I(k) - \nu(c(k) - kI(k)) \\ - [e'(k) + \nu'(c(k) - kI(k))]I(k)H(k)] - bI(k)\}f(k)dk. \tag{20}$$

For any $k \leq \hat{k}$, it follows that $I(k) = 0$. Now, if $I(k) = 1$, for some $k > \hat{k}$, the principal will choose a so as to maximize $a - \nu(a) - \nu'(a)H(k)$. Then condition (19), in conjunction with $H(\cdot)$ being decreasing, implies that the optimal solution in these cases entails maximum effort by the manager, i.e., $a(k) \equiv \bar{a}$. Hence the second-best cutoff, k^*, is as given in (7). Moreover, since $k^* \geq \hat{k}$, with the latter denoting the first-best cutoff, I find that the firm underinvests for all $k \in (\hat{k},k^*)$.

Given the discrete nature of the investment problem, the optimal decision rule, $I(\tilde{k})$, is discontinuous at k^* and insensitive to marginal changes in $\tilde{k}$ elsewhere. Hence, for all types $k \geq k^*$, there only is an incentive to underreport so as to reduce effort. For this to be ruled out, the manager must be paid the above informational rents. In equilibrium, the cutoff-type k^* will just break even, i.e., $s(k^*) = \nu(\bar{a}) - e(k^*)$ and $c(k^*) = k^* + \bar{a}$. Hence, the required cash flow to be delivered at $\tilde{k} = k^*$ renders it unfeasible for a manager with type $k < k^*$ to mimic type k^* so as to trigger the investment decision.[31]

Finally, I show that the optimal mechanism as characterized by (7) is indeed globally incentive compatible. Following Mirrlees (1986), a mechanism that is locally incentive compatible is also globally incentive compatible if $\partial/\partial k U(\tilde{k},k)$ is increasing in $\tilde{k}$. For the solution derived:

$$\frac{\partial}{\partial k} U(\tilde{k},k) = [e'(k) + \nu'(a(\tilde{k},k))]I(\tilde{k}).$$

The Mirrlees (1986) requirement is thus met, because $\nu''(\cdot) \geq 0$, $a(\tilde{k},k)$ is increasing in $\tilde{k}$ and the second-best investment policy, $I(\cdot)$, is upper-tailed.

Proof of Proposition 1

From (9) we know that the optimal required rate of return is given by:

$$R^* \equiv R(k^*) = \frac{k^*}{b} - 1.$$

To compare this term with the firm's cost of capital, I rewrite the first-order condition for the profitability cutoff k^* in (7) as follows:

---

[31] Note that in a more general model with interior effort choices $a(\tilde{k},k)$ in equilibrium, it would be feasible for a manager of type $k < k^*$ to report $\tilde{k} = k^*$ and subsequently to increase his effort to deliver $c(k^*)$, but that would not be optimal, as it would result in negative utility since: $0 = U(k^*) = s(\tilde{k} = k^*) + e(k^*) - \nu(a(\tilde{k}^*,k^*)) > s(\tilde{k} = k^*) + e(k) - \nu(a(\tilde{k} = k^*,k))$, for any $k < k^*$.

$$r = \frac{1}{b}[k^* + e(k^*) - (e'(k^*) + v')H(k^*)] - 1. \tag{21}$$

Hence, a necessary and sufficient condition for $R^* > r$ is that:

$$\delta(k) \equiv e(k) - (e'(k) + v')H(k) < 0,$$

for $k = k^*$. For the two linear specifications of the $e(\cdot)$-function in (PO) and (REP), we have:

$$\delta(k) = (e_0 + e_1 k) - (e_1 + v')H(k).$$

Part 1

Suppose (PO) holds, i.e., $e_0 > 0$ and $e_1 \rightarrow 0$. Then, $\delta(k) \rightarrow e_0 - v'H(k)$, where v' is bounded above by 1 due to the assumed corner solution for $\bar{a}$. It immediately follows that there is a unique value $\hat{e}_0 > 0$ such that $\delta(k^*) > 0$ (and hence $R^* < r$) for any $e_0 > \hat{e}_0$, and vice versa.

Part 2

Suppose (REP) holds, i.e., $e_0 \rightarrow 0$ and $e_1 > 0$. Hence, $\delta(k) \rightarrow e_1[k - H(k)] - v'H(k)$. Since I have not restricted the initial investment expense b from below, for any v' and any type distribution function $F(k)$ over the support $[0,\bar{k}]$, there exists a value b sufficiently low, such that $k^* \rightarrow \underline{k} = 0$. Hence, $e(k^*) \rightarrow e(\underline{k}) \rightarrow 0$. Since both e_1 and v' are assumed to be strictly positive, it follows that $\delta(k^*) < 0$ and, hence, $R^* > r$.

Proof of Proposition 2

Suppose the principal sets the contract parameters as follows:

$$\beta = v', \tag{22}$$

$$\alpha = v(\bar{a}) - v'\bar{a}, \tag{23}$$

and the capital charge rate, ρ^*, is chosen so that:

$$v'[k^* - (1 + \rho^*)b] + e(k^*) = 0. \tag{24}$$

I now show that, given the contract parameters given in (22)–(24), the manager will (i) take the same actions and (ii) earn the same informational rent as in the centralized second-best benchmark solution. For comparison, the manager's informational rent under the centralized mechanism is:

$$U(k,k) = \int_{\underline{k}}^{k} [e'(u) + v']I(u)du = v'[k - k^*] + e(k) - e(k^*), \text{ for } k \geq k^*, \tag{25}$$

and zero otherwise.

Recall that the manager's objective is, for any (α,β,ρ) and any k, to choose $a(\cdot)$ and $I(\cdot)$ so as to maximize:

$$\alpha + \beta[kI + a - (1 + \rho)bI] - v(a) + e(k)I. \tag{26}$$

Hence, given the parameters in (22)–(24), the manager will indeed set $a(k) \equiv \bar{a}$ and $I(k) = 1$ if, and only if, $k \geq k^*$. Moreover, the manager's resulting date–1 informational rent for any $k \geq k^*$ under this delegated mechanism equals $v'[k - (1 + \rho^*)b] + e(k) = v'[k - k^*] + e(k) - e(k^*)$, which coincides with that under centralization, given in (25).

Finally, to see that the optimal capital charge indeed exceeds the firm's cost of capital, r, rewrite (24) as:

$$\rho^* = \frac{k^*}{b} + \frac{1}{v'}\frac{e(k^*)}{b} - 1.$$

By assumption, $v' < 1$. This completes the proof by a simple comparison with (21), since $e(k) > \delta(k)$.

Proof of Proposition 4

First note that in the centralized contracting setting there is a one-to-one mapping from absolute profitability cutoffs, k^*, to required rates of return, R^*, with higher k^* being equivalent to higher R^*. Hence, this proof is confined to absolute profitability cutoffs, k^*. Rewrite (7) for the case of $e(k,I,m)$:

$$k^*(m) + me(k^*(m)) - [me'(k^*(m)) + v']H(k^*(m)) - (1 + r)b = 0. \tag{27}$$

Applying the implicit function theorem to this necessary first-order condition gives (I henceforth suppress the dependence of $k^*(m)$ on m to save on notation):

$$\begin{aligned}\frac{dk^*}{dm} &= -\frac{e(k^*) - e'(k^*)H(k^*)}{1 + me'(k^*) - me''(k^*)H(k^*) - [me'(k^*) + v']H'(k^*)} \\ &= -\frac{e_0 + e_1[k^* - H(k^*)]}{1 + me_1[1 - H'(k^*)] - v'H'(k^*)},\end{aligned} \tag{28}$$

where I have used that, by Definition 1, $e(k) = e_0 + e_1 k$ holds under both (PO) and (REP). Note that the denominator is always positive, hence $sign\{dk^*/dm\} = -sign\{e_0 + e_1[k^* - H(k^*)]\}$.

Part 1

In the plush office (PO) scenario, $e_0 > 0$ and $e_1 \rightarrow 0$. Hence, $dk^*/dm < 0$.

Part 2

Consider next the reputation (REP) scenario, where $e_0 \rightarrow 0$ and $e_1 > 0$. Then $dk^*/dm > 0$ if $k^* < H(k^*)$, and vice versa. Now, it follows from (27) that for any m, there exists a value b sufficiently small, so that $k^* \rightarrow \underline{k}$. By Definition 1, (REP) implies that $\underline{k} \rightarrow 0$ so that $k^* \rightarrow 0$, yet $H(k^*) > 0$ for $k^* \rightarrow \underline{k}$. Hence, $dk^*/dm > 0$ for b sufficiently small. Since $H'(k) \leq 0$, the argument can be reversed to show that for high b, $dk^*/dm < 0$ will hold.

Part 3

Finally, consider a decentralized firm with the manager making the investment decision while compensated based on residual income with ρ as the capital charge rate. Unlike in parts 1 and 2, $e(k)$ is not restricted to be linear in k.

The manager, compensated based on residual income, will invest if and only if $v'[k - (1 + \rho)b] + me(k) \geq 0$. (Recall that the bonus coefficient must equal v' in the optimal solution, due to the arguments given in the main text.) For such a delegated contract to be optimal, this last weak inequality has to hold with equality for $k = k^*$. Solving for the optimal capital charge rate as a function of m yields:

$$\rho^*(m) = \frac{1}{b}\left[k^*(m) + \frac{1}{\nu'}\, me(k^*(m))\right] - 1.$$

Taking the first-order derivative with respect to m gives:

$$\frac{d\rho^*}{dm} = \frac{1}{\nu' b}\left[\nu'\,\frac{dk^*}{dm} + e(k^*(m)) + me'(k^*(m))\,\frac{dk^*}{dm}\right].$$

Now, using (28) above, I can derive a lower bound on this previous expression by using the facts that $\nu' < 1$, $e''(k) \leq 0$ and $H'(k) < 0$ for all k. Since ν' are b both positive, the right-hand side of the last equation is proportional to (again suppressing the argument m n $k^*(m)$):

$$\begin{aligned}
& e(k^*) - [e(k^*) - e'(k^*)H(k^*)] \\
& \quad \times \left[\frac{1 + me'(k^*) - me''(k^*)H(k^*) - [\nu' + me'(k^*)]H'(k^*)}{\nu' + me'(k^*)}\right]^{-1} \\
& = e(k^*)\left[1 - \left(\frac{1 + me'(k^*) - me''(k^*)H(k^*) - [\nu' + me'(k^*)]H'(k^*)}{\nu' + me'(k^*)}\right)^{-1}\right] \\
& \quad + [e'(k^*)H(k^*)]\left[\frac{1 + me'(k^*) - me''(k^*)H(k^*) - [\nu' + me'(k^*)]H'(k^*)}{\nu' + me'(k^*)}\right]^{-1} \\
& > e(k^*)\left[1 - \left(\frac{1 + me'(k^*) - me''(k^*)H(k^*) - [\nu' + me'(k^*)]H'(k^*)}{\nu' + me'(k^*)}\right)^{-1}\right] \\
& > e(k^*)\left[1 - \left(1 - \frac{me''(k^*)H(k^*) - [\nu' + me'(k^*)]H'(k^*)}{\nu' + me'(k^*)}\right)^{-1}\right]. \\
& > 0.
\end{aligned}$$

This last inequality holds, because $e''(k) \leq 0$ and $H'(k) < 0$ for all k imply that $me''(k^*)H(k^*) + [\nu' + me'(k^*)]H'(k^*) < 0$. As a consequence, the term in square brackets is positive and $\rho^*(m)$ is a monotonically increasing function. This completes the proof of Proposition 4.

APPENDIX B
INTERIOR EFFORT CHOICES

In this appendix, I rephrase the model for the case that the condition in (19) does not hold, with the consequence that the principal may want to implement interior effort choices, $a(k) \in (0,\bar{a})$. It is shown here that the preceding results—in particular, the key results in Propositions 1–3—can be generalized to this setting.

The first-best program (if the principal were to know k and could "force" the manager o choose a certain effort level) calls for the investment to be made whenever the present value of cash and empire returns exceeds b. At the same time, the first-best effort level just equates marginal effort costs with marginal effort returns. That is: $\max_{I(k),a(k)}\{(\gamma[k + e(k)] - b)I(k) + \gamma[a(k) - v(a(k))]\}$, so that:

$$\gamma[\hat{k} + e(\hat{k})] - b = 0 \quad \text{and} \quad 1 - v'(\hat{a}(k)) = 0. \tag{29}$$

Note that effort and investment choices are independent in the first-best solution.

In the second-best solution, the principal does not know k and has to elicit this information from the manager. Following similar steps as in the proof of Lemma 1 then yields the following optimization program:

$$\max_{I(k),a(k)} E_k\{[\gamma(k + e(k) - [e'(k) + v'(a(k))]H(k)) - b]I(k) + \gamma[a(k) - v(a(k))]\}.$$

Assuming unique interior solutions, the resulting first-order conditions are:

$$\gamma[k^* + e(k^*) - [e'(k^*) + v'(a(k^*))]H(k^*)] - b = 0, \tag{30}$$

$$1 - v'(a(k^*)) - v''(a(k^*))H(k^*)I(k^*) = 0. \tag{31}$$

Since $I(k) = 1$ if and only if $k \geq k^*$, the investment and effort choices become intertwined for $k \geq k^*$. In comparison with (29), the second-best solution in (30)–(31) is characterized by both underinvestment and underprovision of effort. The latter arises because higher effort levels can only be induced at the expense of higher informational rents. Note that $a(k \geq k^*) \leq a(k < k^*)$, and $a(k)$ is monotonically increasing for all $k \geq k^*$ since $H'(k) \leq 0$. It is easy to show now that Proposition 1 remains qualitatively unchanged once interior effort choices are allowed for.

Following similar steps as in the proof of Lemma 1, the above mechanism can be shown to be globally incentive-compatible, and it results in date–1 informational rents for the manager of $U(k > k^*) = e(k) - e(k^*) + v(a(k)) - v(a(k^*))$ and $U(k \leq k^*) \equiv 0$.

Since the centralized solution now entails type-dependent effort choices, it is clear that a "pure" delegation mechanism, such as that described in Section IV, can no longer replicate this benchmark performance. Instead, I consider the following message-dependent delegation scheme. The manager reports $\tilde{k}$ and thereby picks a compensation scheme from a menu of contracts all of which are linear in residual income: $s(RI|\tilde{k}) = \alpha(\tilde{k}) + \beta(\tilde{k}) \times RI$. Then the manager makes the investment and effort decisions in a decentralized fashion.

Consider the following contract parameters:

$$\beta(k) = v'(a(k)), \tag{32}$$

$$\alpha(k) = v(a(k)) - v'(a(k))a(k) + A(k)I(k), \tag{33}$$

and a hurdle rate, ρ^*, that is implicitly defined by:

$$v'(a(k^*))[k^* - (1 + \rho^*)b] + e(k^*) = 0, \tag{34}$$

where the newly introduced function in (33) equals:

$$A(k) \equiv v(a(k)) - v(a(k^*)) - v'(a(k))[k - k^*] + \left[\frac{v'(a(k))}{v'(a(k^*))} - 1\right] e(k^*).$$

By (31) and (32), $\beta(k) < 1$ for all $k \geq k^*$, even in the case of interior effort choices.

Using (32)–(34), it is a matter of straightforward algebra to show that, having truthfully reported $\tilde{k} = k$, the manager will indeed make the second-best investment and effort choices, $(I(k), a(k))$. The manager's resulting date-1 utility then equals:

$$\begin{aligned}
&\alpha(k) + \beta(k)[a(k) + kI(k) - (1 + \rho^*)bI(k)] - v(a(k)) + e(k)I(k) \\
&= I(k)\left[A(k) + v'(a(k))\left(k - k^* - \frac{e(k^*)}{v'(a(k^*))}\right) + e(k)\right] \\
&= I(k)[e(k) - e(k^*) + v(a(k)) - v(a(k^*))],
\end{aligned}$$

which coincides with his date–1 informational rent under centralization.

By rewriting (34) and (30) in a fashion similar to (15) and (16), above, one finds that $\rho^* > r$, which mirrors Proposition 2. Moreover, since:

$$\rho^* = \frac{1}{b}\left[k^* + \frac{e(k^*)}{v'(a(k^*))}\right] - 1 > \frac{k^*}{b} - 1 \equiv R^*,$$

the key result in Proposition 3 also generalizes to this more general setting.

REFERENCES

Aggarwal, R., and A. Samwick. 1999. Empire-builders and shirkers: investment, firm performance and managerial incentives. NBER Working Paper #7335. Cambridge, MA: NBER.

Antle, R., and G. Eppen. 1985. Credit rationing and organizational slack in capital budgeting. *Management Science* 31 (2): 163–174.

Arya, A., T. Baldenius, and J. Glover. 1999. Residual income, depreciation, and empire building. Working paper, The Ohio State University.

Baiman, S., and R. Verrecchia. 1995. Earnings and price-based compensation contracts in the presence of discretionary trading and incomplete contacting. *Journal of Accounting and Economics* 20 (1): 93–121.

Baldenius, T. 2002. Delegated investment decisions and private benefits of control. Working paper, Columbia University.

———, and A. Ziv. 2003. Performance evaluation and corporate income taxes in a sequential delegation setting. *Review of Accounting Studies* 8: 283–309.

Bareket, M. 2002. Investment decisions under capital constraints: The role of revenue recognition in performance measurement. Working paper, Duke University.

Bernardo, A., H. Cai, and J. Luo. 2001. Capital budgeting and compensation with asymmetric information and moral hazard. *Journal of Financial Economics* 61: 311–344.

Brealey, R., and S. Myers. 2000. *Principles of Corporate Finance*. Boston, MA: McGraw-Hill.

Burrough, B., and J. Helyar. 1990. *Barbarians at the Gate: The Fall of RJR Nabisco*. New York, NY: Harper & Row.

Christensen, P., G. Feltham, and M. Wu. 2002. "Cost of capital" in residual income for performance evaluation. *The Accounting Review* 77 (1): 1–24.

Donaldson, G. 1984. *Managing Corporate Wealth*. New York, NY: Praeger.

Dutta, S., and S. Reichelstein. 2002. Controlling investment decisions: Depreciation and capital charges. *Review of Accounting Studies* 7: 253–281.

———. 2003. Capital budgeting and managerial compensation: Incentive and retention effects. *The Accounting Review* 78 (1): 71–93.

Gibbons, R., and K. Murphy. 1992. Does executive compensation affect investment? *Journal of Applied Corporate Finance* 5 (2): 99–109.

Harris, M., C. Kriebel, and A. Raviv. 1982. Asymmetric information, incentives, and intrafirm resource allocation. *Management Science* 28 (6): 604–620.

———, and A. Raviv. 1996. The capital budgeting process: Incentives and information. *The Journal of Finance* 51 (4): 1139–1173.

———, and ———. 1998. Capital budgeting and delegation. *Journal of Financial Economics* 50: 259–289.

Hart, O., and J. Moore. 1995. Debt and seniority: An analysis of the role of hard claims in constraining management. *American Economic Review* 85 (3): 567–585.

Hennessy, C., and A. Levy. 2002. A unified model of distorted investment: Theory and evidence. Working paper, University of California, Berkeley.

Holmstrom, B., and J. Ricart i Costa. 1986. Managerial incentives and capital management. *Quarterly Journal of Economics* 101: 835–860.

Jensen, M. 1986. The agency costs of free cash flow: Corporate finance and takeovers. *American Economic Review* 76 (2): 323–329.

Jullien, B. 2000. Participation constraints in adverse selection models. *Journal of Economic Theory* 93 (1): 1–47.

Lambert, R. 1986. Executive effort and selection of risky projects. *RAND Journal of Economics* 17 (1): 77–88.

———. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32: 3–87.

Lewis, T., and D. Sappington. 1989. Countervailing incentives in agency problems. *Journal of Economic Theory* 49 (2): 294–313.

Maggi, G., and A. Rodriguez-Clare. 1995. Costly distortion of information in agency problems. *RAND Journal of Economics* 26 (4): 675–689.

Mirrlees, J. 1986. Optimal taxation. In *Handbook of Mathematical Economics*, Vol. 3, edited by K. Arrow, and M. Intrilligator. Amsterdam, The Netherlands: North-Holland.

Myerson, R. 1979. Incentive compatibility and the bargaining problem. *Econometrica* 47: 61–74.

Poterba, J., and L. Summers. 1995. A CEO survey of U.S. companies' time horizons and hurdle rates. *Sloan Management Review* 37 (1): 43–53.

Reichelstein, S. 1997. Investment decisions and managerial performance evaluation. *Review of Accounting Studies* 2: 157–180.

Rochet, J-C., and L. Stole. 2001. The economics of multidimensional screening. In *Advances in Economic Theory*, edited by M. L. Dewatripont et al. Cambridge, MA: Cambridge University Press (forthcoming).

Rogerson, W. 1997. Inter-temporal cost allocation and managerial investment incentives. *Journal of Political Economy* 105: 770–795.

Young, D., and S. O'Byrne. 2001. *EVA and Value-Based Management*. Boston, MA: McGraw-Hill.

THE ACCOUNTING REVIEW
Vol. 78, No. 4
2003
pp. 931–955

Client Importance, Nonaudit Services, and Abnormal Accruals

Hyeesoo Chung
Purdue University

Sanjay Kallapur
University of California, Irvine
and
Purdue University

ABSTRACT: The economic theory of auditor independence (DeAngelo 1981b) suggests that auditors' incentives to compromise their independence are related to client importance. Using ratios of client fees and of nonaudit fees divided by the audit firm's U.S. revenues or a surrogate for the audit-practice-office revenues as measures of client importance, we investigate their association with Jones-model abnormal accruals. In a sample of 1,871 clients of Big 5 audit firms we do not find a statistically significant association between abnormal accruals and any of the client importance measures. Our theory development also suggests that auditor incentives to compromise independence should increase with the extent of client opportunities and incentives to manage earnings, and decrease with the strength of corporate governance and auditor expertise. We also do not find a statistically significant association between abnormal accruals and client importance in subsets of the samples partitioned by proxies for these factors.

Keywords: *auditor independence; earnings management, client importance; nonaudit services.*

Data Availability: *Data are available from public sources identified in the paper.*

We thank Ashiq Ali, Mark Bagnoli, Sudipta Basu, Kevin Den Adel, Bob Eskew, Bill Felix, Frank Heflin, Chuck Horstmann, Iris Jiang, Zhu Liu, Steven Lustgarten, Mark Penno, Sri Ramamoorti, Jean Rothbarth, Wanda Wallace, Susan Watts, and seminar participants at Purdue University, Southern Methodist University, the Information, Markets, and Organizations Conference at the Harvard Business School, and the 2002 AAA Annual Meeting for many helpful comments on this paper.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted September 2002
Accepted April 2003

I. INTRODUCTION

Regulators (Levitt 1998, 2000), press articles (MacDonald 2001; Liesman et al. 2002),[1] and academic studies (Frankel et al. 2002) have recently linked auditors' provision of nonaudit services to poor earnings, and therefore, audit quality. Several other academic studies (DeFond et al. 2002; Francis and Reynolds 2002; Ashbaugh et al. 2003; Francis and Ke 2003) question this link. However, auditors' incentives to compromise their independence are proxied in different studies by different measures: ratio of client nonaudit to client total fees (DeFond et al. 2002; Frankel et al. 2002; Francis and Ke 2003), client total, audit, and nonaudit fee ranks (Frankel et al. 2002), and total fees (Francis and Reynolds 2002; Ashbaugh et al. 2003).

The economic theory of auditor independence (Watts and Zimmerman 1981; DeAngelo 1981b) suggests that auditors' incentives to compromise their independence are related to client importance, the ratio of quasi rents specific to the client divided by all other quasi rents. Following previous studies (Stice 1991; Lys and Watts 1994) we use the ratio of a client's total fees to audit firm's total revenues as our first proxy for client importance. Because it has been argued that nonaudit services have a higher margin than do audit services (SEC 2000, para. III (C)(2)(a)(i) and footnotes 85 and 87), and therefore generate more quasi rents (revenues minus marginal costs), we also use a second measure of client importance—the ratio of nonaudit fees from a client to the audit firm's total U.S. revenues. In our sample these measures are highly correlated (R = 0.97) with each other and also with total fees used by Ashbaugh et al. (2003) and Francis and Reynolds (2002). The correlation in our sample with Frankel et al.'s (2002) ratio of nonaudit to total fees from a client, however, is much lower (corr(client fee to auditor revenue, fee ratio) = 0.26).

If auditors care more about revenues of their practice office rather than the whole firm, then client importance should be assessed at the local office level (following Reynolds and Francis 2001). So we also use two additional proxies for client importance, client total fees and nonaudit fees divided by a surrogate measure of practice-office revenues. To develop this surrogate measure we assume that clients are audited through the audit firm's practice office closest in distance to the client's headquarters, and (similar to Reynolds and Francis 2001) allocate audit firm revenues to practice offices in proportion to the sum of log(sales) of clients of each office.

Consistent with prior research (Becker et al. 1998; DeFond and Subramanyam 1998; Francis et al. 1999) we use abnormal accruals estimated using the (modified) Jones model (Jones 1991; Dechow et al. 1995) as our measure of earnings, and therefore, audit quality. Our sample consists of all companies that reveal audit and nonaudit fees in proxy statements filed between February 5 and June 30, 2001, have data on Compustat, and meet certain criteria (such as being clients of Big 5 audit firms, and not having changed auditors during 2000, n = 1,871).

After controlling for industry effects and determinants of abnormal accruals based on previous studies, we do not find a significant cross-sectional association between the absolute value of abnormal accruals and any of the client importance ratios. The lack of significance is unlikely to have resulted from low power of the Jones-model abnormal accruals to detect earnings management. Even at the upper limit of the 95 percent confidence interval for the client importance coefficient, the difference in the absolute value of abnormal accruals for firms at the first and third quartiles by client importance is much lower than the difference in abnormal accruals of clients of Big 6 and non-Big 6 auditors found in previous research. We can accordingly reject the hypothesis that the difference in the absolute value of abnormal accruals for firms at the first and third quartiles by client importance is as high as the difference found in previous research between abnormal accruals of clients of Big 6 and non-Big 6 auditors.

---

[1] In a widely quoted speech, then SEC Chairman Levitt (1998) expresses concern about erosion of the quality of accounting earnings, and its ultimate effect on investor trust and market volatility. In another speech, Levitt (2000) links accounting manipulation to the substantial growth in consulting services provided by auditors. An article in *Forbes* (MacDonald 2001) asserts that in five of the greatest audit failures of the past two decades (involving Just For Feet, Oxford Health Plans, Colonial Realty, BCCI, and DeLorean Motor Company), the pursuit of consulting contracts compromised auditors' independence. The *Wall Street Journal* (Liesman et al. 2002) suggests that the dual role of auditor and consultant played by Andersen is one of the most serious causes of the Enron debacle.

We also develop a model based on the economic theory of auditor independence to derive cross-sectional client characteristics that could be associated with greater auditor independence impairment. Our model, which is more elaborate than those in previous studies, suggests that independence impairment could be higher in subsamples of clients that have greater opportunities or incentives to manage earnings, or those with weaker corporate governance structures. In addition, auditor expertise may create *client* dependence on auditor and, thereby, strengthen auditors' independence. We measure earnings management opportunities using business and geographical diversification (as proxied by the entropy of business and geographical segment sales). We use leverage, market-to-book ratios, sales growth, and the number of shares and options owned by executives as a percentage of shares outstanding as proxies for client incentives. The strength of corporate governance is proxied by whether the chairman and CEO positions are separate, the percentage of nonemployee directors, and percentage of shares owned by 5 percent blockholders. Auditor expertise is proxied by whether they are the industry specialist. For each continuous variable we divide the sample into three groups by the magnitude of the variable concerned, e.g., firms with low, mid, and high business segment entropy. For indicator variables such as whether the chairman and CEO are separate or are not, we divide the sample into two groups. We test for an association between client importance ratios and abnormal accruals for the sample partitions, and fail to find a significant association in any subsample.

Our results using client importance as a measure of auditor incentives to compromise independence therefore contradict those of Frankel et al. (2002), who instead use the ratio of client nonaudit to total fees. Our analysis indicates that it is difficult to interpret the nonaudit to total fees ratio within DeAngelo's (1981b) framework as a proxy for auditors' incentives to compromise their independence. However, if the provision of certain specific types of nonaudit services impairs auditors' independence regardless of client importance, then the nonaudit-to-total-fees ratio might proxy for the probability that some such service has been provided. Empirically, we find that Frankel et al.'s (2002) finding of an association between the nonaudit-to-total-fees ratio and absolute abnormal accruals is not robust to the inclusion of industry controls (dummy variables for two-digit SIC codes). Our analysis and tests therefore cast doubt on the validity of Frankel et al.'s (2002) conclusions.

We organize the rest of the paper as follows. In Section II we derive a more elaborate version of DeAngelo's (1981b) model of auditor independence and use it to motivate our choice of auditor independence measures and sample partitions where independence impairment is more likely. We also describe our proxy for audit quality. In Section III we describe our data, and in Section IV, our research design and findings. Results of testing for an association between client importance and absolute abnormal accruals within subsamples are presented in Section V. Section VI analyzes, replicates, and tests the robustness of Frankel et al.'s (2002) findings. Section VII summarizes the paper and provides concluding comments.

II. MEASURES OF AUDITOR INDEPENDENCE AND AUDIT QUALITY

Client Importance as Proxy for Auditor Incentives to Compromise Independence

DeAngelo (1981a) and Watts and Zimmerman (1981, 1986) define auditor independence as the probability that an auditor will report a discovered breach.[2,3] DeAngelo (1981b) shows that an auditor's incentive to compromise independence with respect to a client depends on client importance, i.e., the ratio of quasi rents specific to that client divided by the sum of all other quasi rents. Despite having been derived by DeAngelo (1981b), we again derive this result below for purposes of subsequently deriving cross-sectional client characteristics that could be associated with auditor independence impairment. Because DeAngelo's (1981b) purpose was to show why an *auditor* characteristic, auditor size, is (negatively) associated with independence impairment, her model is not elaborate enough for our purpose.

An audit firm's value consists of the present value of future quasi rents or rents[4] specific to a given client, and that of other future quasi rents.[5] Quasi rents arise because of the nature of auditing technology, specifically the higher costs at the initiation of the audit, which generate costs to client of switching auditors (DeAngelo 1981a). Therefore, auditors with at least some bargaining power over clients can raise their prices above marginal costs in subsequent periods, giving rise to quasi rents. Quasi rents can also arise as a mechanism to bond audit quality (as Klein and Leffler [1981] explain in the more general context of bonding product quality), or as an equilibrium response by clients who wish to influence the auditor to permit the client to pursue its chosen reporting goals (Dye 1991). The audit firm's value, V, can then be written as:

$$V = QR_c + QR_o,$$

where QR_c denotes the present value of future quasi rents specific to the given client, and QR_o denotes the present value of all other future quasi rents.

Assuming that an auditor has discovered a breach by the given client, she faces a choice of reporting or not reporting the breach. If she reports the breach, then she risks being fired by the client, in which case all the client-specific quasi rent stream will be forfeited. Denoting as P_{fire} the probability that the client will fire an auditor for reporting the breach, and assuming that QR_o remains unaffected, the auditor's expected wealth if she reports the breach decreases by $P_{fire}QR_c$ to $(1 - P_{fire})\ QR_c + QR_o$. Alternatively, if the auditor refrains from reporting the breach, then she will retain the client, and therefore the client-specific quasi rents. However, there is a risk of being detected—either by analysts in the investment community or the press, by regulators, or *ex post* because of restatements by the client or worsening of the client's financial condition possibly leading to bankruptcy. Detection might

---

[2] This contrasts with the AICPA (GAAS) that defines independence as a mental attitude. Further, some distinguish between independence in fact and independence in appearance. The economic theory and our empirical tests based thereon deal with independence in fact.

[3] In our empirical estimation we are unable to distinguish between breaches that the auditor detected and waived versus those that the auditor did not detect. However, to the extent that detection is a function of auditors' choice variables such as audit effort, the theory applies to such breaches also.

[4] Quasi rents are *ex post* rents after costs have been sunk, i.e., revenues minus marginal costs *excluding* the sunk costs. Rents are *ex ante*, i.e., before costs have been sunk, and therefore equal revenues minus marginal costs *including* those yet to be sunk. Marginal costs in each case include opportunity costs. The distinction between rents and quasi rents is relevant only for whether auditors have any monopoly pricing power; it is therefore irrelevant here.

[5] Some of the quasi rents derived from others will be specific to those respective clients, but others need not be client-specific. For example, if a firm derives quasi rents from past investments in industry expertise, then such quasi rents are not specific to any given client, but are instead industry-specific.

entail loss of reputation, in turn leading to being fired by some clients (DeAngelo 1981b, 190). Another penalty might be lower fees from new clients who subsequently hire the audit firm (Davis and Simon 1992). A wealth-maximizing auditor will refrain from reporting the breach (i.e., compromise independence) if the benefits outweigh the costs, i.e., denoting the probability of detection as P_{detect}, and the fraction of other quasi rents lost upon detection as α, if:

$$QR_c + (1 - P_{detect}\,\alpha)QR_o > (1 - P_{fire})\,QR_c + QR_o,$$

i.e.,

$$QR_c/QR_o > P_{detect}\,\alpha/P_{fire}.$$

This result has the following implications. First, if the ratio $P_{detect}\,\alpha/P_{fire}$ is high enough, then the auditor will not compromise independence for any client. Thus, whether auditors compromise independence or do not is an empirical question. Second, if the auditor compromises independence, then it is when the present value of quasi rents specific to the given client is high in relation to the present value of all other quasi rents. In other words, the incentives to compromise independence depend on client importance—the ratio of the present value of quasi rents from that client to that of all other quasi rents. Third, if P_{detect} depends on the extent of manipulation involved, more manipulation will be permitted for clients from whom the auditor receives higher quasi rents.

Empirically, the quasi rent ratio is unobservable. DeAngelo (1981b) argues that the ratio of fees from a client divided by the audit firm's total revenues is a good proxy for the quasi rent ratio, and previous studies such as Stice (1991) and Lys and Watts (1994) use this ratio to proxy for client importance. Thus, our first proxy for auditor incentives to compromise independence is the ratio of client fees (audit + nonaudit fees paid by the client to the auditor) to the total U.S. revenues of the audit firm. We obtain information on client fees from proxy statements, and on the total U.S. revenues of audit firms from the annual data on the top 100 CPA firms published in the *Accounting Today* magazine (and its website http://www.electronicaccountant.com) in 2001. Second, it has been argued that the nonaudit services have a higher margin (and therefore higher quasi rents, which are defined as revenues minus marginal costs), and that firms underbid on audit services in order to gain access to the nonaudit business (SEC 2000, para. III (C)(2)(a)(i) and footnotes 85 and 87). If so, then the quasi rents could depend more on nonaudit fees rather than total fees. We therefore also use the ratio of nonaudit fees from a client to the total U.S. revenues of the audit firm[6] as another proxy for client importance.

Reynolds and Francis (2001) argue that client importance is more appropriately assessed at the local practice office rather than at the firm-wide level, i.e., as client fees divided by revenues from other clients of the same audit office. Our model implies this under two possibilities: if auditors care more about revenues of their local practice office rather than those of the whole firm, or if the office responsible for a failed audit suffers more than the rest of the firm upon detection. As for the first possibility, Reynolds and Francis (2001) point out that the individual office is the decision-making unit within the audit firm wherein auditors contract with clients, administer audits, and issue audit reports. Thus, if auditors maximize quasi rents of their local practice office rather than of the audit firm overall, then the proper denominator for the client importance ratio is quasi rents from clients of the same audit office. The second possibility is consistent with the results of Chaney and Philipich (2002, Table 5) who find that when Andersen disclosed shredding of Enron documents, its clients suffered a 2 percent decline in stock price, but the clients of the Houston office (which conducted the Enron audit and shredded documents) suffered a 4 percent decline.

---

[6] We are unable to use the auditor's total *nonaudit* revenues as the denominator because this information is not available on the same basis as the proxy disclosures: revenue data disclosed to the SEC Practice Section of the AICPA and in the *Accounting Today* surveys are broken down into those from audit and assurance, tax, and management advisory services. Any audit or assurance services (e.g., WebTrust attest services or audit of the pension plan) not directly related to the audit of the financial statements, however, are classified as nonaudit services for purposes of the proxy disclosure. If the ratio of nonaudit to audit revenues is similar across the sample firms' auditors, then regression R^2s and t-statistics will be unaffected.

We therefore also examine alternative client importance ratios at the practice-office level—client total and nonaudit fees deflated by a surrogate measure of revenues of the practice office that audited that client. We estimate this surrogate measure by assuming that the audit firm's office closest in distance to client headquarters conducts the audit (see the Appendix for details). Reynolds and Francis (2001) instead determine the identity of each client's audit office manually from the 10-K filings by clients.[7] As for revenues by practice office, both Reynolds and Francis (2001) and we estimate it by allocating audit firm revenues to audit offices in proportion to the sum of log(sales) of each office's clients.

Manual data collection for a random sample of 25 clients confirms the accuracy of distance-based matching: 19 of the 25 audit offices, i.e., over 75 percent, are correctly matched. Five of the six mismatches are between audit offices located close to each other, for example San Francisco and San Jose, California, or Washington, D.C. and MacLean, Virginia. By successfully replicating Reynolds and Francis's (2001) results (see the Appendix for details) we confirm that the mismatches do not affect the overall assessment of office-level client importance—possibly because each such office is big (i.e., has many clients), and the mismatches tend to offset each other (i.e., one client wrongly matched to San Francisco instead of San Jose offsets the effect of another client wrongly matched to San Jose instead of San Francisco). In one of the 25 cases the match was wrong: Secure Computing (NASDAQ:SCUR) is headquartered in San Jose but audited by Ernst & Young's office in Minneapolis where SCUR has administrative and research offices.

All of the above proxies at the audit-firm as well as practice-office levels assume that future quasi rents are proportional to current revenues. Previous research (DeBerg et al. 1991) suggests that auditors earn lower consulting revenues from new clients. Because the assumption that future revenues are proportional to current revenues may not be descriptive of such clients, in our estimation we exclude all firms that had new auditors for the sample year. Also, the Big 5 audit firms earn premium fees (Craswell et al. 1995) and have higher internal standards than professional standards require. This is consistent with their suffering a higher penalty for detection of a breach. We control for this difference between the Big 5 and other audit firms by restricting our sample to clients of the Big 5 auditors (there are only 119 observations with non-Big 5 auditors).

Cross-Sectional Differences in Auditor Incentives to Compromise Independence

The parameters α, P_{detect}, and P_{fire} could vary cross-sectionally. For example, Reynolds and Francis (2001) argue that auditors suffer higher penalties for audit failures in large clients (i.e., α is higher for large clients). Audit failures in large clients also might have higher probability of being detected (i.e., higher P_{detect}). Higher α and P_{detect} reduce auditor ncentives to compromise independence. In addition, different groups of clients might have different propensities to fire an auditor for reporting the breach. The higher the P_{fire}, the greater are the auditor's incentives to compromise independence. Below we enumerate the factors that could cause such cross-sectional variation, and explain how we partition our sample to control for the variation.

---

[7] By matching clients with audit offices based on distance we are able to save the effort of collecting this data manually (in 2000 the Big 5 audit firms had almost 13,000 SEC registrants as clients).

If P_{detect} or α is higher for large clients, then it is possible that the quasi rent ratio exceeds $P_{detect}\ \alpha / P_{fire}$ for clients in certain size groups but not for others. The effect of size s ambiguous because larger firms are more important to the auditor, but the probability of detection of breach by such clients and its adverse effect on the auditor are also higher. We present additional results using separate piecewise-linear estimation for three different size groups to address this issue.

P_{detect} could also differ between groups of clients that have different levels of opportunity to manage earnings—the difficulty of detection creates client opportunities for earnngs management, makes it difficult for the auditor to detect it, as well as creates auditor ncentives to overlook it. As explained more fully below, we use business and geographical segment diversification to proxy for client opportunities to manage earnings.

P_{fire} is higher for clients with strong incentives to manage earnings—they are more likely to fire auditors for reporting a breach. We use leverage, market-to-book ratios, sales growth, and employee stock options as a percentage of shares outstanding to control for client incentives to manage earnings. P_{fire} could also differ based on the strength of corporate governance—a more independent board of directors could prevent earnings manipulation n the first place and thereby avoid having to fire the auditor because there is no breach for the auditor to report. Such boards could also protect auditors *ex post*, preventing managers from firing auditors. We use the separation of CEO and chairman positions, percentage of nonemployee directors, and percentage of shares held by 5 percent blockholders as proxies for the strength of corporate governance. Finally, P_{fire} could be lower if the client depends on the auditor for expertise, as Goldman and Barlev (1974) argue. We use auditor ndustry specialization to proxy for auditor expertise. We justify the proxies below and describe how we measure them.

Business and Geographical Segment Diversification

Combining diverse operations gives rise to information asymmetries (Habib et al. 1997), which in turn could give management an opportunity to manage earnings. It could also suppress the activity of information intermediaries (Gilson et al. 2001). Consistent with prior studies (Palepu 1985; Barton 2001) we use entropy as a measure of business or geographical diversification. We obtain data on sales of operating or geographical segments from the Compustat Segment file.

Leverage

DeFond and Jiambalvo (1994) find that the Jones-model abnormal accruals are positive for firms about to violate covenants, and Press and Weintrop (1990) find that leverage is associated with the probability of violating covenants. Thus, high-leverage (measured as debt-to-book equity) firms are more likely to manage earnings.

Market-to-Book Ratio and Sales Growth

Skinner and Sloan (2002) show that the consequences of not meeting earnings forecasts are severe for growth firms. Thus, growth firms, proxied by high market-to-book ratio, and by sales growth, are *a priori* more likely to engage in earnings management. Given the benefits of a smooth pattern of earnings, they have incentives to increase earnings during bad times, and decrease earnings during good times.

Stock Options

The use of stock options as compensation provides incentives to manage earnings (D'Avolio et al. 2001; Byrne 2002). We use the ratio of stock and stock options held by managers divided by total shares outstanding to proxy for such incentives (data source: Standard & Poor's ExecuComp).

Corporate Governance Proxies

Hermalin and Weisbach (2003) argue that effective board monitoring (as reflected in the quality of their decisions on firing CEOs for poor performance, responding to hostile takeovers, adopting poison pills, and design of CEO compensation schemes) is more likely when insiders have less influence on the board. Our proxies for insider influence on the board of directors are (1) separation of CEO and chairman positions (following, among other studies, Tsui et al. [2001], who link it to audit risk); and (2) the percentage of nonemployee directors (following several previous studies, including Weisbach [1988] showing that it is related to the quality of board decisions); and (3) percentage of shares owned by 5 percent blockholders (following Dechow et al. [1996], who link it to financial statement frauds). We obtain data on the above variables from Disclosure GlobalAccess.

In addition to the above studies linking board independence to audit risk and financial fraud, previous studies have found that board independence also proxies for the independence and activity of audit committees, which are responsible for determining audit fee and independence arrangements. Collier and Gregory (1999) find that audit committee activity is lower in U.K. firms when the CEO is the chairman. Menon and Williams (1994) and Beasley and Salterio (2001) find that audit committee independence and activity is positively associated with the percentage of nonemployee directors for U.S. and Canadian firms, respectively.

Auditor Industry Specialization

Clients of industry-specialist auditors might be less likely to fire them because it would entail losing access to the auditor's expertise. We designate the audit firm with the largest number of clients in each two-digit SIC code as the industry specialist, consistent with Craswell et al. (1995).[8]

There are, however, other cross-sectional differences for which we are unable to control. For example, it is possible that not all nonaudit services have the effect of compromising independence. ASR No. 264 mentions tax advice as an example of a service that has obvious economies of scope with the audit function (for auditing the deferred tax accrual). If there are economies of scope and some of the resulting cost savings are passed on to the client, then a client would have to pay more to purchase such services from a party other than the auditor; a threat to stop purchasing such services is therefore less credible than a threat to stop purchasing other kinds of services. Also, an expectation of recurring purchase of a service creates more incentives for the auditor than does a one-time purchase. However, the proxy disclosures do not separate the different kinds of services so we are unable to control for the possibility that the inclusion of fees for such services in the ratios weakens the power of our tests.

---

[8] Craswell et al. (1995) also determine the auditors' market share weighted by sales. Our results using the alternative method are similar.

Abnormal Accruals as Proxy for Audit Quality

Following previous studies (Becker et al. 1998; DeFond and Subramanyam 1998; Francis et al. 1999) we use abnormal accruals estimated using the cross-sectional modified Jones model (Jones 1991; DeFond and Jiambalvo 1994; Dechow et al. 1995) as our measure of earnings management. More recent research is also consistent with our research design choice. Nelson et al. (2002) find that the most frequently attempted method of earnings management by managers in their sample involves reserves (and 35 percent of these attempts are prevented by auditors, Nelson et al. [2002, Table 2]), and these are reflected in accruals.[9] Henninger (2001) finds the Jones-model abnormal accruals to be associated with litigation against auditors. Bartov et al. (2001) show that the cross-sectional Jones-model abnormal accruals are associated with audit qualifications.

We estimate the following model:

$$ACC = a_0\, 1/TA_{-1} + a_1\, (\Delta SALES - \Delta AR)/TA_{-1} + a_2\, PPE/TA_{-1} + e, \qquad (1)$$

where:

ACC = total accruals deflated by TA_{-1}; following Hribar and Collins (2002), we define accruals as the difference between earnings (before extraordinary items and discontinued operations) and cash from operations;
TA_{-1} = total assets at the beginning of the year;
$SALES$ = sales;
AR = accounts receivable; and
PPE = property, plant, and equipment.

Following previous studies (DeFond and Jiambalvo 1994; Bartov et al. 2001), we estimate the model separately for each industry (two-digit SIC code). We calculate abnormal accruals, *DACC*, as the residual from the above model. For each industry, we estimate Model (1) using all firms belonging to that industry (including sample and nonsample firms) with data in Compustat for the year 2000. Because Levitt (1998) expresses concerns about "cookie-jar accounting," implying that firms manage earnings downward as well as upward, we use the absolute value of abnormal accruals in subsequent tests. Becker et al. (1998), however, argue that audit firms are never sued for permitting accruals that decrease earnings, and accordingly signed abnormal accruals are the correct measure. We also report the results of tests using this alternative measure.

The Jones-model abnormal accruals are affected by some possibly nondiscretionary items such as gains on sale of assets. To control for the effect of such items, we present results of an alternative estimation excluding special items (following Xie 2001).

III. DATA AND DESCRIPTIVE STATISTICS

We collect from proxy statements information on audit, IT, and other fees paid by companies to their auditors. Disclosure was required in proxy statements beginning February 5, 2001. Because the SEC Final Rule (SEC 2000) requires proxy statements to disclose "financial information systems design and implementation fees," we searched the 10-K Wizard (http://www.10kwizard.com) for proxy statements filed between February 5 and June 30, 2001 and containing the text in quotes above. We match the resulting 3,965 hits with Compustat, which leaves us with 3,046 firms (see Table 1). From these we delete financial firms because discretionary accruals might not be meaningful for them (n = 778), firms that changed auditors during 2000 (n = 169), firms that did not have a Big 5 auditor (n = 119), and firms that did not have data required for estimating the dependent and independent variables (n = 109). We can thus use 1,871 observations in tests of client importance at the audit firm level. Zip codes of client firm headquarters, required for distance-based matching with audit offices, are unavailable for another 93 firms (in most cases because the headquarters are in foreign countries, e.g., Bermuda), so we have a sample of 1,778 observations for tests of client importance at the local-audit-office level.

---

[9] Under Statement on Auditing Standards No. 57, *Auditing Accounting Estimates* (para. .04), the auditor is responsible for evaluating the reasonableness of accounting estimates made by management.

Our 1,871 sample firms belong to 54 different industries (two-digit SIC codes). Table 2 presents descriptive statistics for the 1,871 sample firms on the variables used in subsequent regressions. The mean and median values of |*DACC*|, the absolute value of discretionary accruals calculated as residuals from Equation (1), are 22.2 and 8.3 percent of total assets, respectively. Mean and median values of the other variables are generally similar to those in Frankel et al.'s (2002) sample. The median client accounts for only 0.01 percent of the revenues of the audit firm (*client/rev*), even the third quartile is only 0.03 percent. The mean is higher, 0.046 percent, suggesting that the ratio is skewed to the right.[10] The median client accounts for 0.6 percent of the revenues of the office through which it is audited, which is almost 60 times as high as the client importance ratio at the audit firm level. The mean, 3.3 percent, is proportionately even higher (3.3/0.046 = 72 times as high as mean client/rev at the audit firm level).

Correlations (not reported) indicate that client importance estimated using total (*client/rev*) and nonaudit fees (*non_aud/rev*) are highly correlated (R = 0.97). *Client/rev* is also highly correlated (R = 0.94) with total fees. The correlation between client importance at the audit firm (*client/rev*) and local office levels (*client/officerev*) is 0.70. Correlations between Frankel et al.'s (2002) fee ratio (nonaudit to total fees) and client importance

TABLE 1
Sample Selection Criteria

Selection Criteria	Observations
Proxy statements filed between 2/5/01 and 6/30/01 with fee data available	3,965
Less: Observations not available on Compustat	(919)
	3,046
Less: Financial institutions (SIC codes 6000–6999)	(778)
Firms that changed auditors during 2000	(169)
Firms with non-Big 5 auditors	(119)
Firms with missing data required for estimating dependent and independent variables	(109)
Sample observations for tests of client importance at the audit firm level	1,871
Less: Observations in which zip codes of client firm headquarters are unavailable on Compustat Research Insight	(93)
Sample observations for tests of client importance at local office level	1,778

---

[10] The non-normality does not affect conclusions from our subsequent regressions; all results are unchanged when we use normal scores of the client importance measures instead of the untransformed variables.

special items (0.6 * Compustat item 17, following Xie [2001]) that could be potentially nondiscretionary.

Accepting the Null Hypothesis

An insignificant coefficient implies an inability to reject the null hypothesis of no association between client importance and the absolute value of abnormal accruals, either because of no association, or because of noise resulting in low power of the tests. The latter possibility, in turn, could be attributable to noisy client importance measures or low power of the Jones-model abnormal accruals to detect differences in audit quality. We follow suggestions in Greenwald (1975) to argue below against the noise explanation and in favor of accepting the null hypothesis.

Our client importance measures are based on theory and have been found to be associated with litigation against auditors (Stice 1991; Lys and Watts 1994); they therefore are unlikely to be too noisy. The Jones model also is unlikely to lack power to detect differences in audit quality—previous studies (Becker et al. 1998; Francis et al. 1999) are able to detect significant differences between abnormal accruals of firms audited by Big 6 versus non-Big 6 auditors.

Even at the upper 95 percent confidence limit, the difference in absolute abnormal accruals between firms whose *client/rev* fall at the third and first quartiles respectively, is small. It is 0.28 percent of total assets,[12] which is lower than the difference between abnormal accruals of clients of Big 6 and non-Big 6 auditors in previous studies. Francis et al. (1999, Table 3) find a 3.6 percent difference (= 0.173 − 0.137) between absolute abnormal accruals of firms audited by non-Big 6 and Big 6 auditors. However, they do not control for other factors that might cause differences in absolute abnormal accruals of clients of non-Big 6 and Big 6 auditors. Becker et al. (1998) do control for other factors and find a 1.5 percent difference; however, their dependent variable is abnormal accruals, not absolute abnormal accruals.

Alternatively, instead of using interval estimation, we can argue in favor of the null in a hypothesis-testing framework as follows. Frame a new null hypothesis as H_0: after controlling for other factors, the difference in $|DACC|$ between firms at the first and third quartiles of *client/rev* is 1.5 percent or larger (the lower of the differences between abnormal accruals of clients of non-Big 6 and Big 6 firms in previous research). We can reject this new null hypothesis comfortably ($t = -8.97$).[13]

V. RESULTS OF PARTITIONING THE SAMPLE

As described in Section II, auditor independence impairment is more likely in certain subsets of client firms. To determine whether the client importance ratios (we describe results for *client/rev*, but they are similar for the other client importance measures) are associated with $|DACC|$ in partitions of the sample, we form dummy variables representing the partition. We interact the client importance ratio with the dummy variables for the partitions. We also include all but one of the dummies as independent variables in the regression (if we included all the dummies, then their sum would be collinear with the sum of the industry dummies). Thus, to investigate the association between |*DACC*| and *client/rev* for firms partitioned into three groups by size (total assets), we estimate the following regression:

---

[12] Our coefficient on *client/rev* (−2.668) and its associated standard error (7.00 = 2.668/0.381) imply that the upper 95 percent confidence limit for the coefficient is 11.052 (= −2.668 + 7.00*1.96). The inter-quartile range for *client/rev* is 0.000254 (= third quartile, 0.000303 − first quartile, 0.000049, from Table 2). Hence, the upper 95 percent confidence limit for the difference in the absolute value of abnormal accruals between firms whose *client/rev* fall at the third and first quartiles respectively is 0.28 percent (= 11.052*0.000254) of total assets. Although the above discussion pertains to the coefficient on *client/rev* in Table 3, the points are equally applicable to coefficients on the other client importance variables in Table 3; we omit those calculations for brevity.

[13] Since our coefficient is −2.668, it implies that the estimate of difference between firms at the 25th and 75th percentiles is $-2.668 * 0.000254 = -0.0007$. The t-statistic for the difference between −0.0007 and 0.015 is $(-0.0007 - 0.015)/(7.00 * 0.000254) = -8.97$.

$$
\begin{aligned}
|DACC| = {} & \textstyle\sum_j a_j D_j + b_1 \log(TA) + b_2 OCF + b_3 OCF^{+} + b_4 ACC_{-1} + b_5 ACC^{+}_{-1} \\
& + b_6 ROA_{-1} + b_7 ROA^{+}_{-1} + b_8 ACQ + b_9 ISSUE + b_{10} TENURE \\
& + b_{11} I_{size=small} + b_{12} I_{size=mid} + b_{13} client/rev \times I_{size=small} \\
& + b_{14} client/rev \times I_{size=mid} + b_{15} client/rev \times I_{size=large} + u, \qquad (3)
\end{aligned}
$$

where:

$I_{size=small}$ = 1 if the firm falls in the smallest group when the sample is divided into 3 groups by size, and 0 otherwise;
$I_{size=mid}$ = 1 if the firm falls in the middle group when the sample is divided into 3 groups by size, and 0 otherwise; and
$I_{size=large}$ = 1 if the firm falls in the largest group when the sample is divided into 3 groups by size, and 0 otherwise.

Coefficients b_{13}, b_{14}, and b_{15} measure the association between *client/rev* and |*DACC*| for groups of small-, mid-, and large-sized clients, respectively, and are reported in the respective columns of Table 4. Recall from Section II that the other partitioning variables are business and geographic segment diversification as measures of opportunity to manage earnings; leverage, market-to-book ratios, sales growth, and use of employee stock options as measures of incentives to manage earnings; separation of CEO and chairman's positions, percentage of nonemployee directors, and percentage of shares owned by 5 percent blockholders as measures of corporate governance; and auditor industry specialization as a measure of client dependence on auditor. The coefficient on *client/rev* is not statistically significant in any sample partition (see Table 4). Moreover, the coefficients also imply small differences between |*DACC*| of firms within the subgroups. The only exception is the coefficient for small firms (+298.706), which implies that the difference in |*DACC*| between small firms in the third and first quartiles by *client/rev* is 1.6 percent of total assets (the interquartile range of small firms' *client/rev*, not reported in any table, is 0.000055). However, given the statistical insignificance of the coefficient, we cannot draw conclusions about auditor independence impairment within this subgroup.[14] Because data on the partitioning variable is unavailable for some sample firms, the number of observations differs for the different partitioning variables.

Our regression specification constrains the coefficients on the control variables (but not the intercepts) to be equal for the sample partitions. We also relax this constraint by estimating the regressions separately for the partitions. The results are similar; for brevity, we do not present the results of separate estimation. We performed factor analysis to determine whether any common factor explains variation in most of the partitioning variables. The

---

[14] Recall that we have no prediction about whether there would be greater auditor independence impairment in large or small firms. The t-statistic reported in the table, 1.14 (corresponding to a two-tailed p-value of 0.25) therefore *overstates* the true significance level—the true p-value is the probability that the largest of three subgroup t-values exceeds 1.14, which is higher than 0.25.

TABLE 4
Parameter Estimates of the Coefficient on Client Importance for Sample Partitions

Firms grouped by size, business segment entropy, geographical segment entropy, leverage, market-to-book ratio, sales growth, percentage of shares and options owned by managers, whether the CEO is the chairman of the board of directors, percentage of outside directors, percentage of shares owned by 5 percent blockholders, and whether the auditor is the industry specialist.

Panel A: Size	**Small, n = 616**	**Mid, n = 620**	**Large, n = 617**
client/rev coefficient	298.706	−45.669	−5.030
t-statistic	(1.140)	(0.501)	(−0.796)

Panel B: Proxies for Opportunities to Manage Earnings

Business Segment Entropy	**Low, n = 852**	**Mid, n = 137**	**High, n = 490**
client/rev coefficient	−9.660	32.522	0.080
t-statistic	(−0.513)	(0.871)	(0.010)
Geographical Segment Entropy	**Low, n = 519**	**Mid, n = 369**	**High, n = 436**
client/rev coefficient	−20.487	0.460	−1.406
t-statistic	(−0.925)	(0.057)	(−0.167)

Panel C: Proxies for Client Incentives

Leverage	**Low, n = 614**	**Mid, n = 621**	**High, n = 620**
client/rev coefficient	−48.631	−3.106	0.562
t-statistic	(−1.217)	(−0.311)	(0.075)
Market-to-Book Ratio	**Low, n = 620**	**Mid, n = 614**	**High, n = 616**
client/rev coefficient	1.161	−6.131	1.561
t-statistic	(0.079)	(−0.552)	(0.206)
Sales Growth	**Low, n = 614**	**Mid, n = 612**	**High, n = 605**
client/rev coefficient	0.174	−0.216	−6.106
t-statistic	(0.025)	(−0.017)	(−0.401)
Stock and Stock Options Owned by Executives as a Fraction of Shares Outstanding	**Low, n = 233**	**Mid, n = 236**	**High, n = 237**
client/rev coefficient	2.058	−4.493	−1.259
t-statistic	(0.477)	(−0.478)	(−0.085)

Panel D: Corporate Governance Factors

Is CEO the Chairman?	**No, n = 707**	**Yes, n = 1108**	
client/rev coefficient	−3.993	1.230	
t-statistic	(−0.375)	(0.164)	
Percentage of Outside Directors	**Low, n = 636**	**Mid, n = 626**	**High, n = 550**
client/rev coefficient	−7.911	−2.666	3.590
t-statistic	(−0.588)	(−0.239)	(0.434)
Percentage of Shares Owned by 5% Block-holders	**Low, n = 599**	**Mid, n = 604**	**High, n = 602**
client/rev coefficient	2.175	−2.275	3.765
t-statistic	(0.313)	(−0.159)	(0.116)

(*continued on next page*)

TABLE 4 (continued)

Panel E: Proxies for Client Dependence on Auditor

Is Auditor the Industry Specialist?	No, n = 1233	Yes, n = 622
client/rev coefficient	−5.638	2.927
t-statistic	(−0.647)	(0.364)

***, **, and * denote significance at the 1, 5, and 10 percent (two-tailed) levels, respectively.
The table reports coefficient values and t-statistics for the interactions of *client/rev* with the partitioning variables. With size, for example, the coefficients reported are those for b_{13}, b_{14}, and b_{15} in the following regression:

$$\begin{aligned}|DACC| = {} & \textstyle\sum_j a_j D_j + b_1 \log(TA) + b_2 OCF + b_3 OCF^{+} + b_4 ACC_{-1} + b_5 ACC^{\pm}_{-1} + b_6 ROA_{-1} + b_7 ROA^{\pm}_{-1} \\ & + b_8 ACQ + b_9 ISSUE + b_{10} TENURE + b_{11} I_{size=small} + b_{12} I_{size=mid} + b_{13} client/rev \times I_{size=small} \\ & + b_{14} client/rev \times I_{size=mid} + b_{15} client/rev \times I_{size=large} + u,\end{aligned}$$

$I_{size=small}$ = 1 if firm falls in the smallest group when the sample is divided into three groups by size, and 0 otherwise;
$I_{size=mid}$ = 1 if firm falls in the middle group when the sample is divided into three groups by size, and 0 otherwise;
$I_{size=large}$ = 1 if firm falls in the highest group when the sample is divided into three groups by size, and 0 otherwise;
size = total assets;
business segments entropy = $\Sigma_{i=1}^{n} P_i \ln(1/P_i)$ where n is a firm's number of business segments and P_i is the share of the *i*th segment in the total sales of the firm; geographical segments entropy is similarly estimated;
leverage = total liabilities divided by stockholders' equity; and
sales growth = $(sales_t/sales_{t-1}) - 1$.

See Table 2 for definitions of other variables.
Except for grouping by indicator variables such as whether the CEO is the chairman, firms are divided into equal-sized groups. Where the groups are not of equal size, e.g., grouping by business or geographical segment entropy, it is because firms with the same value of the partitioning variable are always assigned to the same group.

largest common factor, however, explains only 16 percent of the variation, and the top four factors, only 51 percent. Therefore, we did not perform an analysis using these factors instead of the original partitioning variables. Finally, we attempted to test for independence in the most impaired group, as determined by all factors combined. There is not a single firm in such a group (a firm with high business and geographic segment entropy, leverage, market-to-book ratio, sales growth, and stock options; low percentage of outside directors, and stock ownership by blockholders; whose CEO is chairman and which is audited by a non-industry-specialist auditor), so we could not perform such a test.

VI. NONAUDIT TO TOTAL FEE RATIO AS AN AUDITOR INDEPENDENCE MEASURE

The Earnscliffe Research and Communications (1999) survey finds that there is a perception of auditor independence impairment when the amount of nonaudit fees is large *in relation to audit fees*. This measure has also been used in previous research by Glezen and Millar (1985), Parkash and Venable (1993), Firth (1997), and Frankel et al. (2002). Use of the ratio as a measure of auditor independence impairment would imply that, e.g., the extent of Arthur Andersen's independence impairment with respect to UBICS Inc (an IT services company), which paid audit fees of $45,000 and nonaudit fees of $47,155, is approximately the same as that with respect to Enron, which paid audit and nonaudit fees of $25 million and $27 million, respectively, which is unlikely. Burton (1980) also forcefully argues that the amount of nonaudit fees in relation to audit fees is perceived to impair independence only by those who do not understand the process of auditing (and as mentioned before, our paper examines independence in fact rather than independence in appearance).

However, because Frankel et al. (2002) find a positive relationship between this measure and client importance and interpret their results to mean that higher consulting fees impair auditor independence, we discuss and replicate their finding and examine its robustness below. We are unable to reconcile the fee ratio with the theory of auditor independence developed in Section II. Empirically, Frankel et al.'s (2002) findings are not robust to controls for industry effects.

It is difficult to interpret the nonaudit-to-total-client-fee ratio as being related to auditors' incentives to compromise their independence under the DeAngelo (1981b) framework. Because the ratio is calculated using audit and nonaudit fees at the individual client level, and for many large audits the partner-in-charge is likely to be exclusive to that client,[15] it is tempting to view the ratio as a proxy for incentives at the individual-auditor level. Supporting this view, it has been reported (Brown and Dugan 2002) that beginning in 1998 Andersen had internal targets requiring partners to bring in twice as much consulting revenue as audit revenue.[16] The problem with this interpretation, however, is that it would be hard to make the case that one partner who is in charge of a client that pays less consulting fees in relation to audit fees has lower incentives to appease and keep the client than another partner who is in charge of the audit of a client that pays more consulting fees, given that both are the sole clients of the respective partners-in-charge. Besides, risk assessments are periodically carried out at the national-office level (Beneish et al. 2001), and it is not clear why the national office considers the ratio of nonaudit to audit fees, and not absolute amounts of fees involved.

One interpretation that could support the use of the nonaudit to audit fee ratio as a measure of auditor incentives to compromise independence is as follows. It could be that there are only certain specific types of nonaudit services that impair an auditor's independence—this is the logic underpinning regulatory and legislative actions to ban specific nonaudit services as do the SEC Final Rule (SEC 2000) and the Sarbanes-Oxley Act of 2002. As the ratio of nonaudit to audit fees increases, it could become more likely that the nonaudit services provided include one of those that impair independence.

Empirically, Frankel et al.'s (2002) findings are not robust to controls for industry effects. Using a regression model similar to model (2), but replacing the *client/rev* ratio with the ratio of client nonaudit fees divided by client total fees, and excluding fixed-industry effects as do Frankel et al. (2002), we are able to replicate their findings—we obtain a 0.11 coefficient on the fee ratio variable (t-statistic = 3.17, see Table 5).[17] When we introduce industry (two-digit SIC codes) dummies, the t-statistic drops substantially (to 1.52), which is insignificant at the 5 percent level. Further analysis (not tabulated) shows

---

[15] For example, Enron was the sole client of David Duncan, Andersen's partner-in-charge on the Enron audit (Brown and Dugan 2002).

[16] The interpretation of audit fees here, however, is as a measure of the potential amount of consulting fees possible. Interpreted this way, auditors should be willing to compromise independence for those clients whose nonaudit fees are less than twice their audit fees. This suggests a negative relationship between the nonaudit to audit fee ratio and earnings management, at least when the ratio is less than 2.

[17] There are other differences between our study and Frankel et al. (2002)—we estimate the Jones-model separately for each industry, and our sample excludes clients of non-Big 5 auditors and clients that changed auditors in 2000. Given that we can replicate Frankel et al (2002), these differences do not seem to affect the results.

TABLE 5

Parameter Estimates for Regressions of |*DACC*| on Control Variables and Fee Ratio, with and without Industry Controls

Regression Model 1:

$$|DACC| = b_0 + b_1 \log(TA) + b_2 OCF + b_3 OCF^{+} + b_4 ACC_{-1} + b_5 ACC^{+}_{-1} + b_6 ROA_{-1} + b_7 ROA^{+}_{-1} + b_8 ACQ + b_9 ISSUE + b_{10} TENURE + b_{11} non/tot + u$$

Regression Model 2:

$$|DACC| = \sum_j a_j D_j + b_1 \log(TA) + b_2 OCF + b_3 OCF^{+} + b_4 ACC_{-1} + b_5 ACC^{+}_{-1} + b_6 ROA_{-1} + b_7 ROA^{+}_{-1} + b_8 ACQ + b_9 ISSUE + b_{10} TENURE + b_{11} non/tot + u$$

Control Variables	**Model 1 (without industry controls)**	**Model 2 (with industry controls)**
Intercept	0.123***	
log(*TA*)	−0.014***	−0.004
OCF	−0.190***	−0.192***
OCF^{+}	0.491***	0.535***
ACC_{-1}	0.000	0.000
$ACC_{-1}{}^{+}$	0.048	0.053
ROA_{-1}	0.000	0.000
$ROA_{-1}{}^{+}$	−0.030	−0.065
ACQ	0.038***	0.004
ISSUE	0.098***	0.074***
TENURE	−0.001	−0.002*
Fee Ratio		
Non/tot	0.107*** (3.168)	0.047 (1.521)
Adj. R^2	0.158	0.385
n	1853	1853

***, **, and * denote significance at the 1, 5, and 10 percent (two-tailed) levels respectively.

We do not report t-statistics on control variables, and report those on the fee ratio in parentheses below the coefficient value.

D_j = 1 if firm belongs to *j*th industry, and 0 otherwise; and

non/tot = ratio of nonaudit to total (i.e., audit + nonaudit) fees.

See Table 2 for definitions of other variables.

that three industries, SIC codes 34 (manufacturing: fabricated metal products), 36 (manufacturing: electrical and electronic equipment), and 73 (business services including advertising, software, data processing, etc.), have high |*DACC*| as well as high nonaudit-to-total-fee ratios. When we drop the firms in these industries, the t-statistic on the fee ratio (without industry controls) drops to 1.41, which is similar to the results in the specification where we control for industry fixed effects. Moreover, the coefficient on the fee ratio is largest for small firms, when we divide the sample into three equal-sized groups according to total assets. The coefficient estimates (t-statistics) for small, medium, and large firms are 0.08 (1.72), 0.03 (0.65), and 0.01 (0.21), respectively. If independence impairment is a function of this ratio, then it is more likely to occur with the largest firms where fee amounts are higher, rather than the smallest firms.

Frankel et al. (2002) also present results using percentile ranks (within auditor) of audit and nonaudit fees as client importance measures. They find that the rank of nonaudit fees is significantly positively associated with $|DACC|$. We can replicate this result using their method. However, when we control for industry fixed effects, the coefficient loses its significance (t = 0.70).

VII. CONCLUSIONS

In this paper we use the economic theory of auditor independence to motivate our choice of client importance measures at the audit-firm and local-office levels to proxy for auditor incentives to compromise their independence. Specifically, our measures of client importance are the ratios of total fees (audit and nonaudit) from a client and of nonaudit fees from the client divided by revenues of the audit firm or the practice office. Following previous studies (Becker et al. 1998; Francis et al. 1999) we use absolute Jones-model abnormal accruals of audit clients as measures of earnings and, therefore, audit quality. After controlling for industry and determinants of abnormal accruals based on previous studies, we do not find a significant cross-sectional relationship between the absolute value of abnormal accruals and any of the client importance ratios. Our findings of lack of significance are unlikely to be attributable to low power of the tests—we are able to reject the hypothesis that the difference in abnormal accruals between firms at the first and third quartiles by client importance is as high as the difference found in previous studies between abnormal accruals of clients of non-Big 6 and Big 6 auditors.

Our model based on the theory of auditor independence also suggests sample partitions in which auditors' independence impairment is more likely. We do not find a significant association between the client importance ratios and abnormal accruals in partitions of the sample based on size, client opportunities (as proxied by business and geographical segment diversification) and incentives to manage earnings (as proxied by leverage, market-to-book ratio, sales growth, and use of stock options), strength of corporate governance (as proxied by separation of CEO and chairman positions, the percentage of outside directors, and percentage of shares held by 5 percent blockholders), and auditor expertise (as proxied by whether the auditor is the industry specialist).

Overall, we do not find evidence consistent with auditor independence impairment as a function of different client fee ratios, either in the full sample or even in the subsamples in which, the theory suggests, auditors' incentives to compromise independence should be greatest. However, it is possible that auditor dependence is significantly related to abnormal accruals in other cases that we have not considered, and our results should be interpreted subject to this caveat. Specifically, data limitations prevent us from distinguishing between recurring and nonrecurring nonaudit services, and between services that enjoy significant economies of scope with the audit function, and those that do not. Besides, we and the other related studies (DeFond et al. 2002; Francis and Reynolds 2002; Frankel et al. 2002; Ashbaugh et al. 2003; Francis and Ke 2003) investigate auditor incentives to compromise independence based on financial interest, not other factors such as personal relationships.

APPENDIX

Estimating Client Importance at the Local-Office Level

To estimate client importance at the practice-office level we need (1) the identity of the specific office that conducted the audit of each client, and (2) total auditor revenues by office. Reynolds and Francis (2001) obtain the identity of each client's audit office from the 10-K filings by clients. Because of the effort required in collecting this data (the Big 5 audit firms have almost 13,000 SEC registrants as clients) we use an approximate method whereby we match each client to audit firm office closest in distance to the client headquarters. As for revenues by practice office, both Reynolds and Francis (2001) and we estimate it using client sales, as described below.

To match clients with offices of audit firms, we start with a list of U.S. offices of the Big 5, available from annual reports filed by each firm with the SEC Practice Section of the AICPA for the year 2000. We located 412 offices of the Big 5 audit firms, 57, the least, for KPMG, and 95, the most, for Deloitte & Touche. Reynolds and Francis (2001) have 499 offices in their sample as of 1996, the reduced number is presumably because of Cooper & Lybrand's subsequent merger with Price Waterhouse.[18] We use Compustat to obtain the zip code of the client firm's headquarters. Compustat Research Insight has headquarters-zip-code and sales data on 5,245 clients of the Big 5 audit firms. For each combination of client firm and office of the client's auditor, we calculate the distance between the zip codes associated with the client's head office and audit firm's office. To calculate distance using zip codes, we obtain the latitude and longitude corresponding to the zip codes (with data downloaded from http://www.qlx.com/zipcoder), and then use the Haversine formula (Sinnott 1984) to calculate the distance based on the latitude and longitude. Having assigned clients to audit offices, we sum the logarithm of sales of clients in each office and allocate the auditor's U.S. revenues to offices in proportion to the sum of log sales of clients of each office (similar to Reynolds and Francis 2001). For example, the sum of log(sales) of PricewaterhouseCoopers' (PwC's) Boston office is 245, and that of all its clients with data in Compustat Research Insight is 6,693. Therefore, \$325 million (= 8,878 * 245/6,693) of PwC's \$8.878 billion total revenues are allocated to the Boston office.

We validate our distance-based matching procedure by replicating Reynolds and Francis's (2001) results. The mean and median office-level influence ratios (ratio of log client sales to sum of log sales of all clients audited by the corresponding office) are 0.076 and 0.044, respectively; these values are virtually identical to 0.076 and 0.039 in Reynolds and Francis's (2001) sample. In addition, when we replicate Reynolds and Francis's (2001) main results by regressing $|DACC|$ on office-level influence and operating cash flows, log sales, leverage, and Altman's (1983) Z-score ($n = 4{,}540$), we obtain a negative and significant coefficient ($t = -4.42$) on influence, similar to their -3.82 t-statistic. These results enhance our confidence in the validity of matching clients with audit offices based on distance.

REFERENCES

Altman, E. 1983. *Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy*. New York, NY: John Wiley & Sons.

Ashbaugh, H., R. LaFond, and B. Mayhew. 2003. Do nonaudit services compromise auditor independence? Further evidence. *The Accounting Review* 78 (July): 611–639.

Barton, J. 2001. Does the use of financial derivatives affect earnings management decisions? *The Accounting Review* 76 (1): 1–26.

Bartov, E., F. A. Gul, and J. S. L. Tsui. 2001. Discretionary-accruals models and audit qualifications. *Journal of Accounting and Economics* 30 (3): 421–452.

---

[18] Both Reynolds and Francis (2001) and we regard each city as a single office whenever an audit firm has more than one office in the same city.

Beasley, M., and S. Salterio. 2001. The relationship between board characteristics and voluntary improvements in audit committee composition and experience. *Contemporary Accounting Research* 18 (4): 539–570.

Becker, C., M. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (1): 1–24.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics.* New York, NY: John Wiley & Sons.

Beneish, M., P. Hopkins, and I. Jansen. 2001. Do auditor resignations convey private information about continuing audit clients? Working paper, Indiana University.

Brown, K., and I. J. Dugan. 2002. Sad account: Andersen's fall from grace is a sad tale of greed and miscues. *Wall Street Journal* (June 7): A1.

Burton, J. 1980. A critical look at professionalism and scope of services. *Journal of Accountancy* 149 (4): 48–56.

Byrne, J. A. 2002. How to fix corporate governance. *BusinessWeek* (5/6/02): 69–78.

Chaney, P., and K. Philipich. 2002. Shredded reputation: The cost of audit failure. *Journal of Accounting Research* 40 (4): 1221–1245.

Collier, P., and A. Gregory. 1999. Audit committee activity and agency costs. *Journal of Accounting and Public Policy* 18 (4–5): 311–332.

Craswell, A. T., J. R. Francis, and S. L. Taylor. 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting & Economics* 20 (3): 297–322.

Davis, L., and D. T. Simon. 1992. The impact of SEC disciplinary actions on audit fees. *Auditing: A Journal of Practice & Theory* 11 (1): 58–68.

D'Avolio, G., E. Gildor, and A. Shleifer. 2001. Technology, information production, and market efficiency. Working paper, Harvard University.

DeAngelo, L. E. 1981a. Auditor independence, "lowballing," and disclosure regulation. *Journal of Accounting and Economics* 3 (2): 113–127.

———. 1981b. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (3): 183–199.

DeBerg, C. L., S. E. Kaplan, and K. Pany. 1991. An examination of some relationships between nonaudit services and auditor change. *Accounting Horizons* 5 (1): 17–28.

Dechow, P., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, ———, and ———. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (1): 1–36.

DeFond, M. L., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17 (1–2): 145–176.

———, and K. R. Subramanyam. 1998. Auditor changes and discretionary accruals. *Journal of Accounting and Economics* 25 (1): 35–67.

———, K. Raghunandan, and K. R. Subramanyam. 2002. Do nonaudit service fees impair auditor independence? Evidence from going concern audit opinions. *Journal of Accounting Research* 40 (4): 1247–1274.

Dye, R. 1991. Informationally motivated auditor replacement. *Journal of Accounting and Economics* 14 (4): 347–374.

Earnscliffe Research and Communications (Earnscliffe). 1999. *Report to the United States Independence Standards Board: Research into Perceptions of Auditor Independence and Objectivity.* New York, NY: Independence Standards Board. Available at: http://www.cpaindependence.org/uploac/Research%2C+Auditor+Independence+and+Objec.doc.

Firth, M. 1997. The provision of nonaudit services by accounting firms to their audit clients. *Contemporary Accounting Research* 14 (2): 1–21.

Francis, J., E. Maydew, and H. C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18 (2): 17–34.

———, and J. K. Reynolds. 2002. Professional service fees and auditor objectivity. Working paper, University of Missouri.

———, and B. Ke. 2003. Do fees paid to auditors increase a company's likelihood of meeting analysts' earnings forecasts? Working paper, University of Missouri.

Frankel, R., M. Johnson, and K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* 77 (Supplement): 71–105.

Gilson, S., P. Healy, C. Noe, and K. Palepu. 2001. Analyst specialization and conglomerate breakups. *Journal of Accounting Research* 39 (3): 565–582.

Glezen, G. W., and J. A. Millar. 1985. An empirical investigation of stockholder reaction to disclosures required by ASR No. 250. *Journal of Accounting Research* 23 (2): 859–870.

Goldman, A., and B. Barlev. 1974. The auditor-firm conflict of interests: Its implications for independence. *The Accounting Review* 49 (3): 707–718.

Greenwald, A.G. 1975. Consequences of prejudice against the null hypothesis. *Psychological Bulletin* 82 (1): 1–20.

Habib, M., D. B. Johnsen, and N. Y. Naik. 1997. Spinoffs and information. *Journal of Financial Intermediation* 6 (2): 153–176.

Henninger, W. 2001. The association between auditor litigation and abnormal accruals. *The Accounting Review* 76 (1): 111–126.

Hermalin, B., and M. Weisbach. 2003. Boards of directors as an endogenously determined institution: A survey of the economic literature. *Economic Policy Review* 9 (1): 7–26.

Hribar, P., and D. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (1): 105–134.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.

Klein, B., and K. Leffler. 1981. The role of market forces in assuring contractual performance. *Journal of Political Economy* 89 (4): 615–641.

Kothari, S. P., A. Leone, and C. Wasley. 2004. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* (forthcoming).

Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, New York, NY, September 28.

———. 2000. Renewing the covenant with investors. Remarks delivered at the NYU Center for Law and Business, New York, NY.

Liesman, S., J. Weil, and M. Schroeder. 2002. Dirty books? Accounting debacles spark calls for change: Here's the rundown. *Wall Street Journal* (February 6): A1.

Lys, T., and R. L. Watts. 1994. Lawsuits against auditors. *Journal of Accounting Research* 32 (Supplement): 65–93.

MacDonald, E. 2001. Was that auditor's opinion really independent? *Forbes* (3/19/01): 109–111.

Menon, K., and J. D. Williams. 1994. The use of audit committees for monitoring. *Journal of Accounting and Public Policy* 13 (2): 121–139.

Myers, J., L. Myers, and T. Omer. 2003. Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *The Accounting Review* 78 (July): 779–799.

Nelson, M., J. Elliott, and R. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.

Palepu, K. 1985. Diversification strategy, profit performance, and the entropy measure. *Strategic Management Journal* 6 (3): 239–255.

Parkash, M., and C. F. Venable. 1993. Auditee incentives for auditor independence: The case of nonaudit services. *The Accounting Review* 68 (1): 113–133.

Press, E. G., and J. B. Weintrop. 1990. Accounting-based constraints in public and private debt agreements: Their association with leverage and impact on accounting choice. *Journal of Accounting and Economics* 12 (1–3): 65–95.

Reynolds, J., and J. Francis. 2001. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30 (3): 375–400.

Securities and Exchange Commission (SEC). 2000. *Final Rule: Revision of the Commission's Auditor Independence Requirements.* 17 CFR Parts 210 and 240, File No. S7-13-00. Washington, D.C.: Government Printing Office.

Sinnott, R. W. 1984. The virtues of the Haversine. *Sky and Telescope* 68 (2): 159.

Skinner, D. J., and R. G. Sloan. 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7 (2/3): 289–312.

Stice, J. D. 1991. Using financial and market information to identify pre-engagement factors associated with lawsuits against auditors. *The Accounting Review* 66 (3): 516–533.

Tsui, J. S. L., B. Jaggi, and F. A. Gul. 2001. CEO domination, growth opportunities, and their impact on audit fees. *Journal of Accounting, Auditing & Finance* 16 (3): 189–207.

Watts, R., and J. Zimmerman. 1981. The markets for independence and independent auditors. Working paper, University of Rochester.

———, and ———. 1986. *Positive Accounting Theory.* Englewood Cliffs, NJ: Prentice Hall.

Weisbach, M. 1988. Outside directors and CEO turnover. *Journal of Financial Economics* 20: 431–460.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76 (3): 357–374.

HE ACCOUNTING REVIEW
ol. 78, No. 4
003
p. 957–981

Price versus Non-Price Performance Measures in Optimal CEO Compensation Contracts

John E. Core
Wayne R. Guay
Robert E. Verrecchia
University of Pennsylvania

ABSTRACT: We empirically examine standard agency predictions about how performance measures are optimally weighted to provide CEO incentives. Consistent with prior empirical research, we document that the relative weight on price and non-price performance measures in CEO cash pay is a decreasing function of the relative variances. Agency theory speaks to the weights in total compensation (annual total pay and changes in the CEO's equity portfolio value), however, and we document that very little of CEOs' total incentives come from cash pay. We also document that variation in the relative weight on price and non-price performance measures in CEO total compensation is an increasing function of the relative variances. The conflicting results using total compensation indicate that existing findings on cash pay cannot be interpreted as evidence supporting standard agency predictions. Based on our results, we suggest approaches for future research on performance measure use in CEO total compensation.

Keywords: *agency theory; cash pay; total compensation; equity incentives.*

Data Availability: *All data used in this study are publicly available.*

I. INTRODUCTION

Whether CEO compensation is adequately described by standard agency models is the subject of an active debate in the academic literature (e.g., Prendergast 2002). An extensive body of theory-based research in accounting uses variants of olmstrom's (1979) agency model to examine the use of performance measures in optimal

---

Ve appreciate helpful comments from Frances Ayres, Sudipta Basu, Larry Brown, Gus DeFranco, Thomas Lopez, . P. Kothari, Rick Lambert, David Larcker, Nandu Nayar, Madhav Rajan, Shiva Rajgopal, Richard Taffler, Bente illadsen, Jerry Zimmerman, two anonymous referees, and seminar participants at Cranfield School of Management, olumbia University, Baruch College of the City of New York, Duke University, Georgia State University, the niversity of Oklahoma, and The University of North Carolina. We gratefully acknowledge the financial support f the Wharton School. This manuscript is a substantially revised version of an earlier working paper titled, "Are erformance Measures Other than Price Important to CEO Incentives?"

ditor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted August 2002
Accepted March 2003

contracts.[1] Two key *ceteris paribus* predictions that derive from this earlier work are: (1) the relative weight on a given performance measure is a decreasing function of the noise in the performance measure, and (2) the relative weight on two performance measures is a decreasing function of the relative noise in the performance measures. When used to predict variation in incentives, these predictions should be applied to CEO *total* compensation, which consists not only of the CEO's annual pay, but also of the change in the value of the CEO's equity portfolio (Baker 1987).

In this paper, we bring together and extend two streams of empirical research on CEO incentives. The first stream relates to empirical work on CEO incentives in economics and finance that focuses on the first of the two predictions described above. This work assumes that firms use stock price as the sole performance measure in CEO contracts, and tests whether the weight on stock price in a CEO's total compensation varies inversely with the noise in the stock price (e.g., Garen 1994; Aggarwal and Samwick 1999; Core and Guay 2002b). This literature considers both the CEO's annual pay and the change in the value of the CEO's equity portfolio when estimating total CEO compensation. The second stream of research, primarily in the accounting literature, recognizes that firms use multiple performance measures, such as stock price and accounting earnings, in CEO contracts. These studies explore the second prediction above and test whether the relative weight on two performance measures varies inversely with noise in the performance measures. In contrast to the literature in economics and finance, however, the accounting literature on relative performance measure weights proxies for total compensation with CEO annual cash pay or annual total pay, and omits the incentive weight that CEOs receive from changes in the value of their equity portfolios.

As examples from the accounting literature, Lambert and Larcker (1987) and Sloan (1993) examine cash pay and find that, consistent with standard predictions, the relative weight on accounting earnings and stock returns is a decreasing function of the relative variances of the two performance measures. More recent research examines directly the details of annual bonus plans. This research finds some evidence to support standard predictions about the weight placed on nonfinancial and individual performance measures (e.g., Bushman et al. 1996; Ittner et al. 1997), and predictions about the choice of performance measures (Murphy 2000).

Annual cash pay regressions, however, omit the incentive weight on stock price provided by noncash annual pay and by the CEO's equity portfolio, and in general cannot be interpreted as evidence of agency predictions.[2] The estimated weights in cash pay are informative about the relative use of price and non-price performance measures in optimal CEO contracts only if: (1) the incentive weights in other annual pay and the CEO's equity portfolio are small compared to the weights in cash pay, or if (2) the relative incentive weights in other annual pay and the CEO's equity portfolio are perfectly correlated with the relative weights in cash pay. We provide evidence that neither of these conditions is descriptive of the data.

For our tests of the agency predictions, we define CEO "total compensation" as the sum of cash pay (salary and bonus), other annual pay such as restricted stock and stock options, and the change in the value of the CEO's stock and option portfolio. Our total compensation measure is similar to those used by Antle and Smith (1986), Jensen and

---

[1] For a survey of this literature, see Lambert (2001). For a recent discussion of predictions that derive from this literature, see Bushman and Smith (2001).

[2] For example, Bushman and Smith (2001, 264) argue that the estimated slope coefficients in cash pay regressions do "not have the theoretical interpretations derived from the model."

Murphy (1990), Hall and Liebman (1998), and Aggarwal and Samwick (1999), which include both the CEO's annual pay and changes in the value of the CEO's portfolio of stock and options.

To test the relative use of price and non-price performance measures in CEO contracts, we derive an expression for the ratio of a CEO's incentives provided by a performance measure orthogonal to stock price relative to the incentives provided by the stock price. Consistent with Lambert and Larcker (1987) and Banker and Datar (1989), we derive this expression within a standard agency model under the assumption that an optimal contract is written over price and another performance measure. This incentive ratio is a transformation of the ratio of the weights on the two performance measures, and has the advantage that it can be estimated without specifying the non-price performance measures used. We show that this ratio of incentives is equivalent to the variance of compensation unexplained by stock price, divided by the variance of compensation explained by stock price.

Using a sample of 8,015 CEO-year observations for 2,271 CEOs during the period from 1993 to 2000, we find that the incentives provided by cash pay are small relative to the incentives provided by total compensation. Our results indicate that for 81 percent (57 percent) of the CEOs in our sample, the incentives provided by annual cash pay (annual total pay) are less than 10 percent of the incentives provided by the CEO's stock and option portfolio. These results suggest that conclusions about the use of non-price performance measures in cash pay or annual total pay can be different from conclusions about the use of non-price performance measures in total compensation.

To explore the possibility of differing conclusions using total compensation, we replicate previous findings using CEO cash pay that find that the relative incentives from two performance measures are *decreasing* functions of the relative variances in the performance measures. Next, using total CEO compensation in place of cash pay, we show that the relative incentives from price and non-price measures are *positively* related to the relative variances of the performance measures. These results show that results using cash pay cannot be interpreted as evidence of standard predictions.

We discuss three explanations for our findings on total compensation. First, observed contracts might not be optimal because, for example, CEOs take positions in their own stock that lead to suboptimal incentives from the firm's perspective. Second, observed contracts may be optimal, but their variation is explained by contracting models that emphasize project selection, and are not well explained by the standard agency model where effort aversion is the primary agency problem. Finally, the unexpected positive relation between the relative incentives and the relative variances may be due to measurement error in the relative variance variable as a proxy for relative noise in the performance measures. In particular, the relative variances of the performance measures may instead proxy for the relative sensitivities of the performance measures to the agent's action. If this explanation is valid, then observed contracts may be optimal and well described by standard agency predictions.

The paper is organized as follows. In the following section, we state our null hypothesis and describe how we measure CEO total compensation. In Section III, we describe our sample selection and variable measurement, and in Section IV, we report our results. We summarize our findings and discuss alternative interpretations and implications for past and future research in Section V. Section VI concludes.

II. HYPOTHESIS DEVELOPMENT: RELATIVE WEIGHTS AND RELATIVE INCENTIVES

Our null hypothesis is that standard predictions in the agency literature apply to CEO compensation. We define these two standard predictions. *Ceteris paribus*, (1) the relative weight on two performance measures is a decreasing function of the relative variances of the performance measures, and (2) the relative weight on a given performance measure is a decreasing function of its variance.

To test the null hypothesis, we derive a measure of the relative weights (and relative incentives) from price and non-price performance measures in total compensation. Since Benston (1985) and Murphy (1985), it is widely recognized that changes in the value of CEO equity holdings constitute a major source of CEO stock price incentives. Subsequent research by Jensen and Murphy (1990), Hall and Liebman (1998), and Murphy (1999) shows that the incentives to increase stock price come primarily from CEO equity portfolios, and not from stock-price-driven variation in the CEO's annual pay. We extend this prior literature by investigating cross-sectional variation in the relative use of price and non-price measures in CEO total compensation.

We follow Banker and Datar (1989) and assume that the firm writes an optimal contract, C, to motivate a risk-averse and effort-averse CEO to take an action, a, to increase the mean of firm value. The firm contracts over price, p, and a second performance measure, y. The two performance measures are noisy signals of the agent's action, and we assume that these measures have a bivariate normal distribution. Under the further assumption that utility is logarithmic, it can be shown that the optimal contract satisfies the form:

$$\frac{1}{U'(C)} = C = \alpha + \beta_p p + \beta_y y \tag{1}$$

where:

$$\beta_p = \frac{\partial E[p]/\partial a - \dfrac{\mathrm{cov}(p,y)}{\mathrm{var}(y)}\,\partial E[y]/\partial a}{k\,\mathrm{var}(p)}, \tag{2}$$

$$\beta_y = \frac{\partial E[y]/\partial a - \dfrac{\mathrm{cov}(p,y)}{\mathrm{var}(p)}\,\partial E[p]/\partial a}{k\,\mathrm{var}(y)}, \tag{3}$$

and $k > 0$ is a constant (e.g., Janakiraman et al. 1992).

In implementing an empirical test of Equations (1) through (3), we cannot observe directly which performance measures receive weight in the contract other than price. That is, we cannot directly observe a firm's choice of y. To address this problem, we follow Janakiraman et al. (1992) and assume that both p and y are valuable for contracting. Under this assumption, y consists of a component, bp, that is perfectly correlated with price (i.e., congruent), and a component, z, that is uncorrelated with price (i.e., orthogonal):

$$y = bp + z. \tag{4}$$

To see the intuition for z, consider a performance measure y, such as earnings, that serves both in valuation (as a determinant of price) and in stewardship (as an indicator of the CEO's performance). Given our assumption that y is valuable for contracting, it contains information z about the manager's action that is orthogonal to the stock price. If price efficiently aggregates value-relevant public information about the firm, then the component z that is orthogonal to price consists of information that is either not relevant for valuation or not public. This information plays a stewardship role because price, while it aggregates information about the firm efficiently, nonetheless aggregates information about the manager inefficiently or incongruently (e.g., Lambert and Larcker 1987; Banker and Datar 1989; Paul 1992; Bushman and Indjejikian 1993; Feltham and Xie 1994).

Substituting this equivalence (and its implication that $\partial E[y]/\partial a = b\partial E[p]/\partial a + \partial E[z]/\partial a$) into Equations (1), (2), and (3) and rearranging terms, we obtain:

$$C = \alpha + (\beta_p + b\beta_y)p + \beta_y z, \tag{5}$$

where:

$$(\beta_p + b\beta_y) = \frac{\partial E[p]/\partial a}{k\ \mathrm{var}(p)} \frac{\mathrm{var}(z)}{\mathrm{var}(y)} \tag{6}$$

and:

$$\beta_y = \frac{\partial E[z]/\partial a}{k\ \mathrm{var}(z)} \frac{\mathrm{var}(z)}{\mathrm{var}(y)}. \tag{7}$$

The incentives provided by a performance measure are equal to the product of the weight on the performance measure in the CEO's contract and the sensitivity of the performance measure to the CEO's actions (i.e., the marginal effect of the CEO's effort on the performance measure). For example, the incentives provided by measures congruent to price are:

$$Incentives_p = (\beta_p + b\beta_y)\partial E[p]/\partial a, \tag{8}$$

and the incentives provided by the measure orthogonal to price are:

$$Incentives_z = \beta_y \partial E[z]/\partial a. \tag{9}$$

From Equations (8) and (9), the relative incentives from price and the measure orthogonal to price are equal to the ratio of the weights placed on the performance measures in the CEO's contract (the first term on the right side of Equation (10)) multiplied by the ratio of the sensitivities of the performance measures to the CEO's actions (the second term on the right side of Equation (10)):

$$\frac{Incentives_z}{Incentives_p} = \frac{\beta_y}{(\beta_p + b\beta_y)} \times \frac{\partial E[z]/\partial a}{\partial E[p]/\partial a}, \tag{10}$$

where:

$$\frac{\beta_y}{(\beta_p + b\beta_y)} = \frac{\partial E[z]/\partial a}{\partial E[p]/\partial a} \times \frac{\mathrm{var}(p)}{\mathrm{var}(z)}. \tag{11}$$

Equations (10) and (11) show that, holding constant the relative sensitivities of the measures to the CEO's actions, both the relative incentives and relative weights are decreasing functions of the relative variance of the measures. In addition, these equations show that, all else equal, the relative incentives and relative weight on a given performance measure decrease as the relative variance of the measure increases.

To estimate Equation (10) empirically, we substitute for the relative sensitivities using Equation (11), and restate the incentive ratio as:

$$\frac{Incentives_z}{Incentives_p} = \frac{\beta_y^2\,\mathrm{var}(z)}{(\beta_p + b\beta_y)^2\,\mathrm{var}(p)}. \tag{12}$$

The denominator is equivalent to the variation in total CEO compensation explained by stock price. The numerator is the variation in total CEO compensation explained by measures orthogonal to price. Within this framework, an empirical estimate of the relative incentives of two performance measures is the ratio of the variation in total compensation explained by the two performance measures.[3]

To estimate the ratios in Equations (10) and (11), we create a broad measure of total CEO compensation that incorporates all identifiable annual changes in CEO firm-specific wealth. Similar to measures used by Antle and Smith (1986), Jensen and Murphy (1990), and Hall and Liebman (1998), we define total compensation as the sum of changes in the value of a CEO's portfolio of stock and options, and annual pay from salary, bonus, grants of options and restricted stock, and other annual pay:

$$\begin{aligned} \textit{Total compensation} &= \textit{Cash Pay} + \textit{Other Annual Pay} \\ &\quad + \textit{Change in Portfolio Value} \\ &= \textit{Total Pay} + \textit{Change in Portfolio Value} \end{aligned} \tag{13}$$

Baker (1987) and Hall and Liebman (1998) argue that to obtain an appropriate measure of monetary incentives, the computation of a manager's compensation should include changes in the value of the CEO's portfolio of stock and options. These authors base this argument on the idea that CEOs are motivated by how their actions change their wealth. The wealth effect of a CEO's action comes not just from annual pay, but also from its effect on his other firm-specific wealth. When choosing his optimal action, the CEO considers not only the weight on non-price and price performance measures in his total pay, but also the weight on price performance provided by his equity holdings.

We note that although our measure of total compensation is a broad measure of the changes in the CEO's wealth, we cannot observe some components of the CEO's wealth. Our incentive measure is incomplete in that it does not include changes in the present value of the CEO's future pay.[4] Second, some of our sample CEOs could be hedging the firm-specific risk in their equity portfolios through derivative securities such as caps and collars that are negatively correlated with firm-specific price changes. These instruments can reduce the CEO's exposure to the firm's stock price and the price-based incentives provided by their portfolios. Recent research by Bettis et al. (2001) indicates that some CEOs use these techniques, but the small sample size in Bettis et al. (2001) suggests that this behavior is limited. Although it is possible that some CEOs engage in this behavior and do not file required SEC disclosures, the fact that secret hedging activities violate SEC disclosure and insider trading rules reduces the likelihood that this behavior is widespread. Finally, a CEO's outside wealth may contain holdings that are negatively correlated with the systematic component of his own firm's performance. For example, CEOs may hedge the systematic component of their firm's stock return by maintaining short positions in competitors' stocks or in market indexes. Although these holdings reduce the CEO's risk, they affect the CEO's incentives only to the extent that the CEO's actions influence the performance of these other firms. Except in highly concentrated industries in which a given CEO's actions affect competitor firms' values, it is unlikely that CEOs' incentives are affected by a correlation between the values of outside equity holdings and firm stock.

---

[3] We thank Rick Lambert for suggesting this equivalence to us.

III. SAMPLE AND VARIABLE MEASUREMENT

We obtain data on CEO option and stock holdings, and option and restricted stock grants, from Standard & Poor's ExecuComp database. We use CRSP data to generate measures of stock returns, stock-return volatility, and treasury bond yields, and we use Compustat as the source for firms' financial data and industry classifications.

Sample Selection

We obtain an initial sample of CEOs from the ExecuComp database. Consistent with Sloan (1993) and Baber et al. (1998), we require that the CEO serve for a full year in the current year, and for a full year in the year prior, which yields an initial sample of 9,161 CEO-year observations from December 1993 through December 2000. We then eliminate 819 observations due to missing CRSP data, 33 observations due to missing Compustat data on earnings, and 272 observations due to missing CEO option portfolio data. We also require that each two-digit SIC code have a minimum of ten observations, which eliminates 22 observations. The final sample consists of 8,015 CEO-year observations from 1993 to 2000.

Variable Measurement

We define "cash pay" as salary plus bonus and obtain these values directly from ExecuComp. We define "total pay" as the sum of cash pay, long-term incentive plan payouts, the value of restricted stock grants, the value of options granted during the year, and any other annual pay. We obtain all of the components of total pay directly from ExecuComp, with the exception of the value of stock option grants. We estimate the grant date value of options granted during the year using a modified version of the Black and Scholes (1973) model.[5] Consistent with the findings of Hemmer et al. (1996) and Huddart and Lang (1996) that employees exercise options prior to maturity, we assume the expected time-to-exercise is 70 percent of the option grant's stated maturity. Our inference, however, is unaffected if we value the grant using the stated time-to-maturity. The quantity of options in the new grant, the grant date stock price, the exercise price, and the time-to-maturity are available from the ExecuComp database. We measure the expected stock-return volatility as the annualized standard deviation of daily stock returns over the 120 trading days preceding the end of the fiscal year in which the grant was made. We estimate expected dividend yield as cash dividends paid in the fiscal year the grant is made divided by year-end stock price. We use the treasury-bond yield corresponding to the option's remaining time-to-maturity to estimate the risk-free rate.

---

[4] The omission of the change in the present value of future pay will likely cause us to understate the incentives from stock price, because the present value of future pay is significantly related to current stock returns (Boschen and Smith 1995; Boschen et al. 2003), but insignificantly related to current accounting performance (Boschen et al. 2003). To the extent that current non-price performance measures affect future turnover probability, we will understate the incentives provided by these measures. Evidence in Jensen and Murphy (1990) suggests that the change in the present value of future pay related to involuntary turnover is small relative to the decline in stock portfolio value that precedes such turnover.

We define total CEO compensation as the sum of total pay and the change in the value of the CEO's stock and option portfolio during the year. We estimate the change in the value of the CEO's stock holdings as the number of shares of stock held by the CEO at the beginning of the year (including restricted stock) multiplied by the change in stock price (including dividends) during the year. Similarly, we define the change in the value of the CEO's option holdings as the number of options held by the CEO at the beginning of the year, multiplied by an estimate of the change in the average option price during the year. By computing value changes in this way, we concentrate on the *ex ante* incentives provided by the CEO's stock and option holdings at the beginning of each year. Thus, we are able to reduce the measurement difficulties caused by the fact that the CEO's stock holdings increase and decrease at various (unobservable) points during the year because of grants of restricted stock, sales and purchases of stock, and exercises of options. We also reduce the measurement difficulties caused by the fact that the CEO's option holdings increase at various (unobservable) points during the year because of option grants and exercises of reload options, and decrease because of option exercises.[6] This method is equivalent to assuming that all grants of restricted stock and options occur at year-end and that all of the CEO's trades and option exercises settle at year-end.

To estimate the value of the CEO's option portfolio at the end of a year *t*, we use the method developed by Core and Guay (2002a). We modify this method by assuming times-to-exercise equal to 70 percent of the Core and Guay assumed times-to-maturity. These authors value the portfolio as if it were three grants of options: the options granted during year *t*, the unexercisable options granted prior to year *t*, and the exercisable options granted prior to year t. To obtain the value of the options held at the end of year $t+1$, we use the same parameters as described above, with the exception that we reduce the stated maturity (time-to-exercise) of each option group by one year (0.7 of one year) and value the portfolio using the stock price at the end of year $t+1$. By so doing, we isolate the variation in the option portfolio value due to price changes, and exclude changes in interest rates, dividend yield, and stock volatility as sources of option portfolio volatility.

---

[5] Some researchers (e.g., Lambert et al. 1991; Hall and Murphy 2002) argue that risk-averse executives attach substantial discounts (over 50 percent in some cases) to the market value (Black-Scholes value), and the change in value, of their stock and option portfolios. The key assumption driving these results is that the length of the single-period agency model is ten years (i.e., there is no annual contracting as is assumed in most empirical agency research). Core and Guay (2001) argue that, under the set of assumptions used in optimal contracting models such as the one examined in this paper, executives attach much smaller discounts (less than 5 percent) to changes in the market value of their stock and option portfolios. Because the maintained hypothesis of this paper is that observed contracts are optimal, we do not incorporate the Hall and Murphy (2002) results into our primary analysis. We show below, however, that our main results are not sensitive to assuming that CEOs attach very large discounts (75 percent) to changes in the market value of their equity portfolios.

[6] One measurement difficulty that remains in our data stems from the fact that firms do not report CEO stock holdings and option holdings as of the same date. Specifically, option holdings are reported in the proxy statement as of the most recent fiscal year-end, whereas stock holdings are usually reported about three months after the end of the most recent fiscal year-end. In cases where the CEO buys stock or exercises stock options and retains the stock during this three-month period, our estimate of stock price incentives will be upward-biased. In cases where the CEO sells stock during this three-month period, our estimate of stock price incentives will be downward-biased. We expect that the overall net effect of this measurement issue is a slight upward bias in our estimate of stock price incentives.

IV. RESULTS

Descriptive Statistics

Table 1 contains descriptive statistics on the CEOs' annual pay and levels and changes in their holdings of stock and options. In Panel A, the mean values of the CEOs' stock and option holdings are \$97.6 million and \$13.2 million, respectively. More than 50 percent of the CEOs hold at least \$13.3 million in total equity holdings and 75 percent of the CEOs hold at least \$5.0 million. Panel B shows that the mean annual salary and bonus is just over \$1.2 million, and on average, comprises about 30 percent of the annual total pay. The

TABLE 1
Descriptive Statistics of CEO Equity Holdings and Total Compensation

The sample consists of 8,015 CEO-years of ExecuComp data (2,271 CEOs) from 1993–2000. Stock holdings is the value of the CEO's common stock, option holdings is the value of the CEO's option portfolio, and total equity holdings is the sum of stock and option holdings. We measure these variables at the beginning of the year in which we compute compensation. Cash pay is salary plus bonus. Other pay is the sum of long-term incentive plan payouts, the value of restricted stock grants, the value of options granted during the year, and any other annual pay. Total pay is the sum of cash pay and other pay. Change in the value of the CEO's stock holdings is the number of shares of stock held by the CEO at the beginning of the year (including restricted stock holdings) multiplied by the change in stock price (including dividends) during the year. The change in the value of the CEO's option holdings is the number of options held by the CEO at the beginning of the year multiplied by an estimate of the change in the average option price during the year. Total compensation is the sum of total pay and the change in the value of the CEO's stock and option holdings.

	Mean	Std. Dev.	Min.	Q1	Median	Q3	Max.
Panel A: Levels of Equity Holdings (\$ millions)							
Equity holdings							
Stock holdings	97.63	1,065.21	0.00	1.61	6.28	23.18	55,987.42
Option holdings	13.22	53.26	0.00	0.91	3.43	9.91	2,309.08
Total equity holdings	110.85	1,072.76	0.00	4.99	13.27	40.50	55,987.42
Panel B: Total Compensation = Total Pay + Change in Value of Total Equity Holdings (\$ millions)							
Annual pay							
Cash pay	1.23	2.11	0.00	0.54	0.87	1.40	102.45
Other pay	2.98	14.77	0.00	0.19	0.75	2.28	793.81
Total pay	4.21	15.77	0.00	0.93	1.76	3.71	883.81
Change in Value of Equity Holdings							
Stock holdings	33.57	767.09	−7,117.53	−0.36	0.22	2.98	39,321.90
Option holdings	1.47	23.26	−436.45	−0.40	0.00	0.65	890.87
Total equity holdings	35.03	773.44	−7,117.53	−1.20	0.35	5.10	39,717.50
Total compensation	39.25	775.56	−7,117.45	0.56	2.52	9.25	39,928.40

remaining 70 percent of annual pay is comprised primarily of stock option grants, restricted stock grants, and long-term incentive payouts. Total equity holdings are large compared to annual pay. Mean and median statistics for the sample CEOs as a group suggest that cross-sectional variation in cash pay and total pay are small relative to cross-sectional variation in the value of CEOs' stock and option portfolios.

The Relative Incentives Provided by Price and Non-Price Performance Measures in CEO Total Compensation

The pooled descriptive statistics reported in Table 1 are influenced by CEOs with extreme total pay or extreme changes in stock and option portfolio value. To compute a more precise measure of the incentives from annual pay relative to total compensation, we group the CEO-year observations in time-series by CEO and construct three variance ratios: annual cash pay variance to the variance of the change in the value of portfolio holdings, annual total pay variance to the variance of the change in the value of portfolio holdings, and the variance ratio defined by Equation (12).

Equation (12) is the ratio of the variance of total compensation explained by all variables other than price to the variance of total compensation explained by price. To estimate this ratio empirically, we define total compensation as in Section II, Equation (13):

$$\textit{Total compensation} = \textit{Total Pay} + \Delta \textit{Port},$$

where *Total Pay* is annual total pay and *ΔPort* is the change in the value of a CEO's portfolio of stock and options. Estimating the relative amount of non-price incentives in total CEO compensation is equivalent to estimating the variation in total compensation that is unrelated to stock returns, relative to the variation that is explained by stock returns. The variation in total compensation unrelated to returns is:

$$\textit{Unexplained variation} = \textit{Var}(\textit{Total compensation} \mid \textit{Returns}). \tag{14}$$

Changes in the value of common stock holdings are completely determined by changes in stock price. Further, performance-related changes in the value of stock option holdings are driven by changes in stock price. Although changes in the other option-pricing inputs (i.e., the risk-free rate, dividend yield, and volatility) do affect the value of an option portfolio, these parameters are unlikely to be performance measures that corporate boards use as alternatives to stock price performance. In our analysis, we minimize the influence of other minor factors on option portfolio value by holding the option-pricing inputs other than price and time-to-maturity constant when estimating the change in option values.

Because all of the variation in the change in equity portfolio value is due to stock returns, the unexplained variation in total compensation may be restated using Equations (13) and (14) as:

$$\textit{Unexplained Variation} = \textit{Var}(\textit{Total Pay} \mid \textit{Returns}). \tag{15}$$

Because changes in the value of a CEO's stock and option portfolio are determined by returns, Equation (15) indicates that performance measures other than the stock price will manifest themselves as determinants of annual pay.

In Table 2, we report results on variance ratios computed separately for each CEO. To obtain reasonable estimates of the amount of pay variation unexplained by returns, we restrict the sample to only the 958 CEOs with at least four annual observations. In Panel A,

TABLE 2
Variance Ratios for CEO Annual Pay and Equity Portfolio Changes

Cash pay is salary plus bonus. Total pay is the sum of cash pay, long-term incentive plan payouts, the value of restricted stock grants, the value of options granted during the year, and any other annual pay. Total compensation is the sum of total pay and the change in the value of the CEO's stock and option holdings. The data presented in the table are descriptive statistics for variance ratios computed for each of the 958 CEOs. We compute each variance for all of the CEO-year observations for that CEO. The first two columns of each panel present the variances of cash and total pay as a fraction of the variance of the changes in portfolio value, where the portfolio is the sum of stock and option holdings. The ratio in the third column of each panel is the ratio of total pay variance that is not explained by returns to the variance of total compensation explained by returns. The unexplained variance in total pay is the total residual variation from by-CEO regressions (minimum of four observations) of the change in total pay from year $t-1$ to year t regressed on an intercept and the stock return during year t. The explained compensation variance is the sum of the variance in the change in the value of the CEO's stock and option holdings plus the variation explained by returns in the by-CEO total pay regressions.

Panel A: Data Aggregated by CEO

The sample consists of 958 CEOs with at least four consecutive annual observations of ExecuComp compensation data from 1993–2000.

	Cash Pay Variance/ Portfolio Variance	Total Pay Variance/ Portfolio Variance	Unexplained Total Pay Variance/Explained Comp Variance
99%	6.12	35.25	36.55
90%	0.12	1.59	1.17
75%	0.02	0.37	0.24
50%	0.00	0.06	0.04
25%	0.00	0.01	0.00
1%	0.00	0.00	0.00

Panel B: Data Aggregated by CEO then by Two-Digit SIC (median CEO values reported for each SIC code)

The sample consists of 958 CEOs, as reported in Panel A. The data are descriptive statistics for variance ratios for the median CEO in each of 54 two-digit SIC codes. We compute these ratios by taking the corresponding ratios for the 958 CEOs reported in Panel A, grouping them by two-digit SIC, and taking the median variance ratio for each two-digit SIC in which there are at least three CEO observations.

	Cash Pay Variance/ Portfolio Variance	Total Pay Variance/ Portfolio Variance	Unexplained Total Pay Variance/Explained Comp Variance
Max	0.04	1.89	1.55
90%	0.02	0.22	0.15
75%	0.00	0.09	0.08
50%	0.00	0.04	0.04
25%	0.00	0.01	0.01
Min	0.00	0.00	0.00

we report results before industry aggregation. In each of the three columns, we rank the CEOs by the computed variance ratio. The first two columns of Panel A of Table 2 summarize the variance ratios before we remove the pay variance explained by returns from the numerator. We interpret these ratios under the strong assumption that *all* variation in annual pay is due to non-price performance measures. The first column shows that the median CEO receives very few incentives from variation in his cash pay relative to the variation in his portfolio, and that cash pay variance is small relative to portfolio variance for most CEOs. The second column shows the median CEO receives modest incentives from variation in his total pay relative to the variation in his portfolio, but that total pay variance provides substantial incentives for many CEOs.

Although this by-CEO procedure has the advantage of partitioning CEO observations into groups with the greatest possible homogeneity, the trade-off is that our estimates of individual total compensation variances are likely to be imprecise because of the small number of time-series observations. Further, the results are likely to be influenced by CEOs who experience unusual *ex post* realizations in their annual pay and changes in stock and option portfolio value over the short time series.[7]

To mitigate the influence of measurement error in our estimates of pay variances over a short time series, we also group the individual CEO variance ratios by primary two-digit SIC code, and report the results in Panel B. To do this, we restrict each SIC group to contain at least three CEO variance-ratio observations as computed in Panel A of Table 2. For each SIC group, we determine the variance ratio of the median CEO. Using the median values reduces the influence of observations with unusual numerator or denominator values on the distribution of variance ratios. The results in Panel B support the inference that the median CEO receives few incentives from either cash or total pay. Some CEOs receive substantial incentives, however, from total pay. The variance in their total pay relative to the portfolio change variance is greater than 9 percent for 25 percent of the sample.

The variance ratios reported in the first two columns of Table 2 likely overstate CEO incentives from non-price performance measures because prior research shows predictable variation in CEO annual pay related to stock price performance. To provide a more precise estimate of the relative use of CEO performance measures other than price, we estimate regressions that remove the portion of annual pay that is attributable to stock price performance. We use the variance of the residuals from these regressions as an estimate of the variation in annual pay that is not due to stock price performance. As shown by Equation (4), any performance measure that is not congruent with stock returns can be decomposed into a portion that is congruent with stock return and a portion that is orthogonal to stock return.

For each CEO, we estimate the following parsimonious model for annual total pay:

$$(Total\ pay_t - Total\ pay_{t-1}) = \beta_0 + \beta_1 Return_t + \varepsilon_t. \tag{16}$$

In this by-CEO changes model, $Total\ pay_{t-1}$ controls for the expected level of the CEO's pay, and $Return_t$ captures variation in the CEO's pay related to stock returns. We measure $Return_t$ as log(1 + annual return). The residual captures the variation in pay due to any non-price performance measure that is orthogonal to price, and any other residual error in the regression model. Because we estimate Equation (16) by CEO with as few as four observations, we do not have enough degrees of freedom to control for other determinants of pay levels, such as the investment opportunity set and firm size. Prior research has shown these firm characteristics to be important determinants of pay levels (e.g., Smith and Watts 1992). Because we estimate Equation (16) by CEO, however, we do not expect our regressions to be influenced by substantial variation in the firm's investment opportunity set or size. Further, excluding these additional explanatory variables reduces the likelihood that we underestimate the variation in pay that is not explained by price.

---

[7] Ideally, we would like to assess the performance measures using *ex ante* weights in CEO contracts. To illustrate the potential problems with using *ex post* realizations, consider two CEOs with beginning stock and option portfolios of equal value, and identical annual pay arrangements based on performance measures other than price. Assume that both CEOs receive the same annual pay changes each year (which are unrelated to price), but for one CEO the stock price has very high realized variation, and for the other CEO the stock price has very low realized variation. Even though the two CEOs have identical contracting arrangements and the same *ex ante* incentives, the computed variance of changes in equity portfolio value (and *ex post* computed incentives) will be very different between the CEOs.

In the third columns of both Panels A and B of Table 2, we create the incentive ratio defined by Equation (12), which is the annual pay variance unexplained by price divided by the total compensation variance explained by price. We measure explained total compensation variance as the variance of the sum of annual pay due to price performance and the change in stock and option portfolio value. The median CEO's ratio of unexplained total compensation variance is less than 0.08 for 75 percent of the SIC groups. Under the assumptions and framework developed in Section II, these findings suggest that for over 75 percent of the CEOs, the incentives provided by stock price are approximately 12 times greater than the incentives provided by performance measures orthogonal to price. These results indicate that stock price performance is the main driver of total incentives for most CEOs. The results also suggest that empirical studies that infer CEO incentives from annual pay are likely to measure total CEO incentives with substantial measurement error.

Empirical Tests of Relative Weights

In this section, we describe our tests of the agency predictions developed in Section II. Following Sloan (1993), we estimate pooled cross-sectional regressions of percentage compensation changes. To compute these compensation changes, we note that current compensation equals expected compensation *plus* unexpected compensation due to performance and other factors. To remove the effect of scale differences in compensation across CEOs, we compute unexpected percentage change in compensation by subtracting expected compensation from current compensation and deflating the difference by expected compensation:

$$\begin{aligned}&\textit{Unexpected percentage change in compensation}\\&\quad = (\textit{Compensation}_t - E_{t-1}[\textit{Compensation}_t])/E_{t-1}[\textit{Compensation}_t] \qquad (17)\\&\quad = (\textit{Compensation}_t/E_{t-1}[\textit{Compensation}_t]) - 1.\end{aligned}$$

Previous work examines the weights on performance measures in cash pay and total pay, and we begin our analysis by examining variation in these components of total compensation. Consistent with the theory and predictions in Section II, we also examine variation in unexpected percentage change in total compensation.

Like Sloan (1993), we regress unexpected percentage change in compensation on a price performance measure, p, and a non-price performance measure, y.[8] To test whether variation in performance measure weights is consistent with the agency predictions, we include two terms that interact the price performance and the non-price performance measures with the ratio of the variances of the non-price performance measure and the price performance measure. As noted above, the prediction of a negative relation between the relative weight on two performance measures and the relative variances relies on an assumption that the relative sensitivities of the performance measures to the manager's actions is held constant. Prior research argues that stock price is more sensitive to the manager's action when growth opportunities are greater, and studies such as Lambert and Larcker (1987) and Baber et al. (1996) provide evidence consistent with this prediction. Following Smith and Watts (1992), we use beginning-of-year book-to-market ratio as a proxy for the inverse of the firm's growth opportunities, and include two terms that interact the price performance and the non-price performance measures with the book-to-market ratio. We estimate three versions of the following specification using cash pay, total pay, and total compensation as the dependent variables:

---

[8] Sloan examines $\log(\textit{Cash pay}_t/\textit{Cash pay}_{t-1})$, which is approximately equal to $(\textit{Cash pay}_t/\textit{Cash pay}_{t-1}) - 1$ when the percentage changes are small. The use of the logarithm helps mitigate the influence of outliers. We do not use the log specification because total compensation can be negative. As described below we mitigate the influence of outliers by winsorizing observations at the 1 percent level.

$$\begin{aligned}&(Compensation_t/E_{t-1}[Compensation_t]) - 1\\&\quad = \beta_0 + \beta_1 p_t + \beta_2 y_t + \beta_3 p_t{*}var(y)/var(p) + \beta_4 y_t{*}var(y)/var(p)\\&\qquad + \beta_5 p_t{*}book\text{-}to\text{-}market_{t-1} + \beta_6 y_t{*}book\text{-}to\text{-}market_{t-1} + u_t.\end{aligned} \tag{18}$$

The key agency predictions we test are that the weights on p and y decrease with the relative noise in each performance measure: that is, β_3 is positive and β_4 is negative. Under the assumption that stock price is relatively more sensitive to the manager's actions when book-to-market is low, we also predict that β_5 is negative and β_6 is positive.

A difficulty in estimating Equation (18) is creating an appropriate estimate of expected total compensation. Unlike cash pay and total pay, lagged total compensation does not provide a reasonable estimate of expected total compensation because it assumes negative values when stock returns are negative (even though expected returns and expected compensation are always positive). To address this issue, we assume that the expected change in (expected return on) the portfolio value is 10 percent of its beginning value. Because options have a higher expected return than common stock, the actual expected portfolio return is a function of the expected risk of the firm's equity and the relative portion of the portfolio value that is options. Our results are not sensitive to using alternative expected return assumptions ranging from 7 percent to 14 percent. Based on this expected return assumption, our estimate of expected total compensation is the sum of lagged total pay and beginning-of-period portfolio value multiplied by 10 percent. Thus, we measure unexpected percentage change in total compensation as $(Total\ pay_t + \Delta Portvalue_t)/(Total\ pay_{t-1} + 0.1 \times Portvalue_{t-1}) - 1$.

We use log(1 + annual return) as our proxy for price performance. Although the standard agency model does not point to specific non-price performance measures, many studies, such as Sloan (1993), use a financial accounting measure of performance as a proxy. Thus, we choose as our measure of accounting performance the change in earnings per share before extraordinary items (*Change in EPS*) scaled by beginning-of-period stock price.

To estimate the variance of earnings changes and the variance of returns, we use a time-series of ten years of returns and earnings changes ending with the last year for which we have compensation data available. We code these variables as missing if there are fewer than eight observations. Because some of our firms have fewer than eight years of data, our sample size is reduced from 8,015 to 6,574.

To mitigate the influence of outliers in the dependent variable and in the variance ratio, we set the lower- and upper-most percentiles of the unexpected percentage changes in cash pay, total pay, and total compensation and the variance ratio equal to the values at the 1st and 99th percentiles, respectively. All of our specifications include an intercept and include year indicator variables (not tabulated).

In Column (1) of Table 3, we confirm prior findings by Lambert and Larcker (1987) and Sloan (1993) that the relative weight on *Change in EPS* and stock returns in CEO cash pay is negatively related to *var(Change in EPS)/var(ret)*. The combination of an insignificant coefficient on the term that interacts returns and *var(Change in EPS)/var(ret)*, and a negative and significant coefficient on the term that interacts *Change in EPS* and *var(Change in EPS)/var(ret)*, supports the standard prediction that the relative weight on

TABLE 3
Regressions of Unexpected Percentage Changes in CEO Cash Pay, CEO Total Pay, and CEO Total Compensation on Hypothesized Determinants

The sample consists of 6,565 CEO-years of ExecuComp compensation data from 1993–2000. Cash pay is salary plus bonus. We use lag cash pay to estimate expected cash pay and compute unexpected percentage change in cash pay as $(Cash\ pay_t/Cash\ pay_{t-1}) - 1$. Total pay is the sum of cash pay, long-term incentive plan payouts, the value of restricted stock grants, the value of options granted during the year, and any other annual pay. We use lag total pay to estimate expected total pay and compute unexpected percentage change in total pay as $(Total\ pay_t/Total\ pay_{t-1}) - 1$. Total compensation is the sum of total pay and the change in the value of the CEO's stock and option holdings. We estimate expected total compensation as the sum of lag total pay plus an assumed 10 percent return on the value of beginning portfolio holdings. We compute unexpected percentage change in total compensation as $(Total\ pay_t + \Delta Portvalue_t)/(Total\ pay_{t-1} + 0.1 \times Portvalue_{t-1}) - 1$. Return is the logarithm of (1 + the continuously compounded monthly return over the firm's fiscal year). Change in EPS is current earnings per share before extraordinary items minus lagged earnings per share divided by the beginning of year stock price. We compute the variance of returns and the variance of change in EPS using a time-series of ten years (minimum of eight observations) of returns and earnings changes ending with the last year for which we have compensation data available. Book-to-market is the ratio of book value of assets to market value of assets measured at the beginning of the fiscal year. The regressions reported in the table also include an intercept and seven year indicator variables (not tabulated). The t-statistics are based on the Huber-White robust standard error, which is a generalization of the White (1980) standard error that is robust to both serial correlation and heteroscedasticity (Rogers 1993).

		Dependent Variable is Unexpected Percentage Change In:		
Independent Variable	**Predicted Sign**	**Cash Pay (1)**	**Total Pay (2)**	**Total Comp (3)**
Return	+	0.21 (7.43)	0.63 (4.39)	5.54 (20.38)
*Return** *var(Change in EPS)/var(ret)*	+	−0.07 (−1.34)	−0.14 (−1.08)	−1.77 (−4.32)
*Return** *Book-to-market*	−	0.09 (1.99)	−0.16 (−0.85)	−1.89 (−4.85)
Change in EPS	+	1.07 (2.76)	1.21 (1.18)	−0.87 (−1.00)
*Change in EPS** *var(Change in EPS)/var(ret)*	−	−0.65 (−4.09)	−0.23 (−1.02)	0.09 (0.26)
*Change in EPS** *Book-to-market*	+	−0.11 (−0.38)	−0.87 (−0.83)	0.99 (1.08)
R^2		14.53%	3.46%	59.48%

Change in EPS and stock returns is negatively related to *var(Change in EPS)/var(ret)*.[9] As do earlier studies, we find that annual cash pay is significantly positively related to both stock return and *Change in EPS*.

Although prior research and our analysis both support the standard prediction with cash pay, previous research has had little success replicating this prediction with annual total pay (e.g., Baber et al. 1998). Column (2) of Table 3 shows that we too find weak (strong) evidence of a significant positive relation between earnings (stock price) performance and unexpected total pay. When we examine total annual pay, however, we find no evidence of a significant relation in the predicted direction between the relative weight on non-price and price performance measures and the variance ratio. The coefficients on the non-price and price performance measures that interact with the noise ratio are insignificant.

These annual pay results are incomplete and potentially misleading because the theory predicting variation in relative weights applies to total compensation, not cash pay or annual total pay (Baker 1987). Because the standard predictions are supported only within cash pay, the results in Columns (1) and (2) of Table 3 uphold the theoretical predictions about the relative weighting of price and non-price performance measures in the optimal contract only if variation in cash pay is highly correlated with variation in total compensation.

In Column (3) of Table 3, we estimate Equation (18) to examine variation in the unexpected percentage change in total compensation. The results in Column (3) show a stark contrast to the standard prediction: the weight on stock returns exhibits a significant negative relation with *var(Change in EPS)/var(ret)*, and the weight on *Change in EPS* shows no significant relation with *var(Change in EPS)/var(ret)*. Thus, in contrast to the standard prediction, the relative weight on price and non-price performance measures in total compensation *increases* with the ratio of the variances.[10]

In addition, in the Column (3) total compensation regression, the coefficient on book-to-market interacted with return (change in EPS) is significantly negative (insignificantly positive). This finding is consistent with earlier findings (e.g., Baber et al. 1996) and our prediction that the relative weight on stock return should increase with growth. We find no significant relation between book-to-market and the relative weight on stock returns in total pay in Column (2). In contrast to our predictions, in the Column (1) cash pay regressions, we find a significant positive relation between book-to-market and the relative weight on stock returns.

In untabulated tests, we find that our Table 3 results are robust to dropping the interactive terms with book-to-market. Specifically, when the book-to-market variables are removed, the relative weight on price and non-price performance measures in total compensation (cash compensation) increases (decreases) with the ratio of the variances. Our results in Table 3 are also robust to replacing *Change in EPS* with a proxy for non-price performance, which we define by z, the component of accounting performance that is orthogonal to stock returns. This approach is consistent with Equation (4). We estimate the variance of z as the variance of the residuals from firm-specific regressions of changes in earnings per share on stock returns (minimum of eight observations).

---

[9] See Baber et al. (1996, 307, footnote 15) for a discussion of this inference.

[10] Sloan (1993) provides other related evidence that the standard predictions are descriptive within cash pay. Specifically, he finds that the relative weight on accounting earnings in cash pay increases with the ratio of the value of CEO stockholdings to cash pay, and the proportion of market-driven variation in stock returns. Consistent with our findings reported in Table 3, Columns (2) and (3), neither of these predictions hold in our sample when tested on total pay or on total compensation.

Empirical Tests of Relative Incentives

Although the total compensation results in Table 3 reject the standard predictions, these findings are not completely satisfying for two reasons. First, in Column (3) of Table 3, we use a previously unexamined measure of unexpected total compensation as the dependent variable. This measure relies on a noisy estimate of expected return on portfolio value (i.e., a constant expected return across all firms). Second, although it is clear that our findings are heavily influenced by CEOs' stock and option holdings, our tests so far do not clearly illustrate *why* these portfolio holdings influence the results.

In Table 4, we provide more evidence of the robustness of the results in Table 3. We examine whether the relative incentives from non-price-based (z) and price-based performance measures in CEO total compensation vary consistently with standard predictions. We first estimate the relation between $var(z)/var(ret)$ and the ratio of incentives. The computation of the ratio of incentives is described in the previous subsection, and the empirical distribution of the ratio is summarized in Column (3) of Panel A of Table 2. As we show in Equation (10), and consistent with the standard prediction about the ratio of the performance measure weights, we predict that the ratio of incentives is negatively related to the variance ratio $var(z)/var(ret)$. As above, to mitigate the influence of outliers in the dependent variable and in the variance estimates, we set the lower- and upper-most percentiles of the dependent variable, the variance ratio, and the variance of returns equal to the values at the 1st and 99th percentiles, respectively. To control for the relative sensitivities of the performance measures to the manager's actions, we include in the regression the average beginning-of-year book-to-market ratio (averaged for the period over which the incentive ratio is computed) as a proxy for the inverse of the firm's growth opportunities. We predict a positive association between this variable and the incentive ratio under the assumption that stock price is more sensitive to the manager's action when growth opportunities are greater.

We present the results of this regression in Column (1) of Table 4. Counter to the standard prediction, we find that the relation between the incentive ratio and the variance ratio is significantly positive. The significant positive relation between book-to-market and the incentive ratio is consistent with our prediction and with previous findings. Both of these findings (a positive association between the variance ratio and the relative weight on stock return and a positive association between growth and the relative weight on stock return) are consistent with our results in Column (3) of Table 3.

A potential shortcoming of the foregoing test is that we must specify a proxy for z to estimate the ratio $var(z)/var(ret)$; this contrasts with estimating the incentive ratio, which did not require us to identify a specific, unobservable performance measure z. Our proxy for z is the component of *Change in EPS* (scaled by price) that is unexplained by stock returns, and is a noisy measure of the non-price performance measure embedded in the incentive ratio. To ensure that our results are not induced by our proxy for z, we test a second standard prediction from Equation (10). Holding the relative sensitivities and the variance of the non-price performance measure z constant, Equation (10) predicts that the ratio of incentives from z and p is an increasing function of the variance of returns (because we predict that less weight is placed on a noisier performance measure). The negative coefficient on $var(ret)$ in Column (2) also rejects this prediction.

Another concern with our calculation of the incentive ratio is that we use the variation in total pay that is orthogonal to stock returns to estimate the incentives due to non-price performance measures. Many non-price performance measures, such as accounting earnings, are known to be correlated with stock returns. Thus, it is possible that by giving

TABLE 4
Regressions of Log(Incentive Ratio) and Log(Total Pay Variance/Portfolio Variance) on Hypothesized Determinants

The sample for Columns (1) through (5) consists of a cross-section of 812 CEO observations. Incentive ratio is the ratio of total pay variance not explained by returns to the variance of total compensation explained by returns. The unexplained variance in total pay is the total residual variation from by-CEO regressions (minimum of four observations) of the change in total pay from year $t-1$ to year t regressed on an intercept and the stock return during year t. Explained compensation variance is the sum of the variance in the change in the value of the CEO's stock and option holdings plus the explained variation in the by-CEO total pay regressions. Total pay is the sum of cash pay, long-term incentive plan payouts, the value of restricted stock grants, the value of options granted during the year, and any other annual pay. z is the residual from firm-specific regressions of changes in EPS on stock returns. Change in EPS is current per share earnings before extraordinary items minus lagged earnings per share divided by the beginning of year stock price. We compute the variance of returns and the variance of z using a time-series of ten years (minimum of eight observations) of returns and earnings changes ending with the last year for which we have compensation data available. Book-to-market is the average beginning-of-year ratio of book value of assets to market value of assets (averaged over the period over which we compute the dependent variable). Market value is market value of assets (averaged over the period over which we compute the dependent variable). In Column (6), we compute a predetermined measure of return variance using a time-series of ten years (minimum of eight observations) of returns ending with the year prior to the first CEO observation. This variable measurement restricts the sample to 623 CEO observations for Column (5). The t-statistics are based on OLS standard errors if the White (1980) test does not indicate heteroscedasticity and on White (1980) standard errors otherwise.

		Dependent Variable				
Independent Variable	**Predicted Sign**	**Log (Incentive ratio) (1)**	**Log (Incentive ratio) (2)**	**Log (Total pay variance/Portfolio Variance) (3)**	**Log (1/Portfolio Variance) (4)**	**Log (1/Portfolio Variance) [restricted] (5)**
var(z)/var(ret)	–	3.31 (3.23)		3.10 (3.05)		
Book-to-market	+	3.21 (5.86)	3.33 (5.25)	3.56 (6.53)	4.76 (9.88)	3.73 (6.39)
var(ret)	+		−0.63 (−2.31)		−1.69 (−5.60)	−3.16 (−6.41)
log(market value)	–				−0.80 (−10.42)	−0.81 (−8.54)
Observations		812	812	812	812	623
R^2		6.82%	6.28%	7.69%	28.42%	22.66%

returns the first opportunity to explain changes in pay, the residuals understate the variation in pay that could be explained by non-price performance measures, and therefore overstate the incentives due to price. To address this concern, we re-estimate the regressions in Columns (1) and (2) using as the dependent variable the ratio of total pay variance to total portfolio variance summarized in Column (2) of Panel A of Table 2. Under the assumption that all of the variation in total pay is due to non-price measures and that only the portfolio provides price incentives, this ratio can be interpreted as a ratio of non-price to price incentives. Column (3) shows that inference from this alternative incentive ratio is unchanged: counter to the standard prediction, *var(z)/var(ret)* has a positive association with the incentive ratio. In untabulated results, consistent with Column (2) and counter to the standard prediction, we also find that *var(ret)* has a negative association with the incentive ratio.[11]

In Column (4), we estimate the relation between the inverse of CEOs' equity portfolio variance and return variance. Unlike the ratio measures in the previous columns, this measure is unscaled, and we include the log of market value of equity to control for the effect of firm size and the level of equity portfolio incentives. Our results indicate that part of the reason our earlier results reject the agency predictions is that there is a strong positive relation between portfolio variance and return variance, which is indicated by the negative relation between the inverse of portfolio variance and return variance. We also find a positive relation between the level of equity portfolio variance and both market value of equity and growth opportunities (for which the inverse of the book-to-market ratio serves as a proxy). This result is consistent with previous findings that equity portfolio incentives increase with firm size and growth.

We estimate *var(Change in EPS)*, *var(z)*, and *var(ret)* over a ten-year period that ends with the last CEO compensation observation. This procedure introduces the possibility of a mechanical relation between the variance and compensation. This concern is particularly important in Column (4) because firms with more variable contemporaneous returns will have more variable portfolio returns. Therefore, we calculate the variances over the period prior to the first CEO compensation observation. This results in a reduction in the sample size. However, Column (5) shows that we obtain the same results on a smaller sample when we examine the determinants of the inverse of CEOs' portfolio variance. We also obtain the same statistical inference from the other results shown in Tables 3 and 4 if we use this modified variance measure (untabulated).

Why Does Total Compensation Not Support the Standard Predictions?

The results in Tables 2, 3, and 4 indicate that: (1) stock price is the primary performance measure in CEOs' incentives, (2) CEOs' portfolio holdings are an extremely important component of stock price incentives, and (3) portfolio holdings do not appear to vary consistent with standard agency predictions.

Although this variation in CEO portfolio holdings is inconsistent with standard agency predictions, it supports prior research findings of a positive association between return variance and CEO ownership incentives (e.g., Demsetz and Lehn 1985; Core and Guay 1999; Bushman et al. 2000; Core and Guay 2002b). An extensive literature predicts and finds that CEO ownership and CEO stock price incentives vary with economic determinants predicted by theory, including firm size, growth opportunities, and return variance (e.g., Demsetz and Lehn 1985; Smith and Watts 1992; Gaver and Gaver 1993; Mehran 1995; Core and Guay 1999; Bushman et al. 2000; Baker and Hall 2002). Our results in Columns (4) and (5) of Table 4 are consistent with the results on the determinants of equity incentives in Demsetz and Lehn (1985), Core and Guay (1999), and Bushman et al. (2000). In particular, our finding that CEO equity incentives are positively related to the variance of stock returns is also shown in this prior research. To summarize, it is difficult to find evidence in CEO total compensation that supports the standard predictions because equity portfolio holdings are large relative to total pay and the incentives provided by these holdings appear to vary contrary to standard predictions.

---

[11] As discussed in footnote 5, researchers such as Lambert et al. (1991) and Hall and Murphy (2002) argue that, under the assumption that CEOs must hold their stock and option portfolios for a ten-year period, CEOs would attach a substantial discount to the value of new equity grants and the change in the market value of their stock and option portfolios. To test the sensitivity of our results, we assume that CEOs value new option grants and their equity portfolios at 25 percent of their market values. Our statistical inference from Table 3, Columns (2) and (3) and Table 4 is unchanged. Thus, our inference does not change substantially when we make the (fairly extreme) assumption that CEOs, because they must hold their portfolios for a ten-year period, only value their portfolios at 25 percent of their market values.

It is worth clarifying how the results in Table 4 are related to the recent debate by Aggarwal and Samwick (1999) and Core and Guay (2002b) about whether stock price incentives in total compensation are consistent with the standard predictions. Aggarwal and Samwick (1999) examine variation in the *level* of total CEO compensation to test an agency model based on Holmstrom and Milgrom (1987) in which firms use a single performance measure in CEO contracts. Aggarwal and Samwick (1999) assume that this performance measure is the dollar change in the firm's market value, and provide evidence that the weight on dollar returns in CEO compensation decreases with the variance of dollar changes in the firm's market value. Core and Guay (2002b) question whether this finding supports their prediction, and show that the variance of dollar change in market value is a better proxy for firm size than it is for performance measure noise. Milbourn (2003, 254–255) estimates regressions that are similar to Core and Guay's (2002b) regressions and finds that the variability of dollar returns "behaves similarly to firm size." Core and Guay (2002b) also show that the weight on stock returns in CEO compensation increases with percentage return variance as an alternative proxy for performance measure noise. Our objective in this paper differs from this prior work in that we test the richer agency model typically examined in the accounting literature in which: (1) firms use multiple performance measures in contracts with CEOs, (2) percentage stock return is the price performance measure, and (3) the weights on the measures are predicted to be a function of the relative variances of the performance measures.

Robustness of Results to Implicit Relative Performance Evaluation in CEO Contracts

Holmstrom (1982) predicts that compensation contracts filter out systematic risk (i.e., relative performance evaluation, RPE). Some argue that because CEO stock and option holdings are exposed to both idiosyncratic and systematic risk, CEO contracts such as those in Table 1 that exhibit large equity holdings are *prima facie* inconsistent with this prediction (e.g., Abowd and Kaplan 1999, 162). As emphasized above, however, contract theory applies to the CEO's total wealth. Broadly stated, if a CEO is endowed with outside wealth that he prefers to hold in the market index, then large firm-specific equity holdings are consistent with an RPE contract. That is, large stock and option holdings are consistent with an RPE prediction that the optimal contract requires the CEO to hold more than his preferred level of firm stock, and less than his preferred level of the market index. To see this, consider a firm that hires a new CEO with $100 in outside wealth that she prefers to hold in the market index. Suppose that the optimal contract with this new CEO requires her to purchase $50 of the firm's stock, which she finances by selling $50 of her market holdings. The contract thus forces her to hold $50 more (less) of the stock (market index) than she prefers. The implicit indexing of her holdings of the firm's stock is not observed because the structure of her outside wealth and her preferences are not observed. A more complete argument is detailed in Core and Guay (2001).

A related argument, discussed above, and summarized in Jin (2002), is that CEOs may hedge the systematic component of their firm's stock return by maintaining short positions in competitors' stocks or in market indices. In other words, the CEO may choose to hold some systematic risk and, to the extent feasible, trade out of any remaining systematic component of the firm's stock. The firm would allow or encourage this behavior because the contract is more efficient when it does not impose more systematic risk than the CEO desires. These arguments suggest that contracts requiring CEOs to hold stock and options impose additional firm-specific risk on the CEO, but do not necessarily impose additional systematic risk on the CEO. If the assumption that the CEO's stock and option portfolio effectively imposes no additional systematic risk is valid, then the appropriate price-based performance measure is the firm's idiosyncratic return, and the standard agency model predicts that the contract weight placed on price performance will be lower when the variance of the idiosyncratic return is higher.

To examine the robustness of our results to implicit RPE, we re-estimate the ratios reported in Table 2 and the regressions reported in Tables 3 and 4 using idiosyncratic returns instead of total firm returns. We estimate idiosyncratic returns as the residuals from a market model using annual firm returns and value-weighted CRSP market returns. As before, we use the ten-year period ending with the last year for which we have compensation data. We then use these idiosyncratic returns to estimate CEO portfolio variances, incentive ratios, percentage change in total compensation, and variance ratios, and re-estimate the results in Tables 2, 3, and 4. The re-estimated Table 2 continues to show that price-based incentives comprise a very large fraction of total incentives for most of the distribution of CEOs. In the re-estimated Table 3, the results for cash pay (total pay and total compensation) continue to be consistent (inconsistent) with agency predictions. In the re-estimated Table 4, the results continue to be inconsistent with agency predictions, and portfolio variance is positively related to both idiosyncratic return variance and firm size, and negatively related to book-to-market. In summary, all of the results are robust to using idiosyncratic firm returns instead of total firm returns.

V. SUMMARY AND INTERPRETATION OF RESULTS

Results above confirm findings in prior research that most variation in CEO total compensation arises from CEO equity holdings. We also show that for a substantial majority of CEOs, annual cash pay variance (salary plus bonus) is less than 1 percent of equity portfolio variance. Further, we find that although variation in the relative weight in annual cash pay supports standard predictions, the variation in the relative weights and the relative incentives in total compensation strongly reject the standard predictions.

Before proceeding, we note that an explanation for our results is that large CEO ownership stakes may occur because CEOs voluntarily own stock for reasons unrelated to optimal contracting. If the CEO chooses to hold a level of stock different from that preferred by the firm, then his incentives may not be consistent with a contract that is optimal from the firm's perspective.

Holding aside the concern that CEO contracts might be simply suboptimal, we now speculate on how to reconcile our results with optimal contracting. One conjecture is that evidence from cash pay should receive greater emphasis than evidence from noncash compensation. The support for this conjecture is that CEOs have a better understanding of how to increase accounting performance measures than stock prices; thus, in practice incentives from cash pay may be stronger than incentives from noncash compensation (e.g., Murphy 1999; Bushman and Smith 2001). If this is true, then it is possible that a modest-sized bonus plan provides nearly all of a CEO's incentives, even in the presence of a very large equity portfolio. Accordingly, support for predictions in cash pay should be given more weight when interpreting the findings because cash pay is a better proxy for total compensation. Theory, however, predicts that efficient contracts will not impose unnecessary compensation risk. Consequently, if the stock price is not sensitive to CEOs' actions, then price-based measures should receive little, if any, weight in CEO contracts, and CEOs with large equity portfolios should not be observed.

Another possibility is that standard predictions describe equilibrium CEO contracts, but, because of a recent time-series shift in the percentage of annual pay delivered in the form of option grants, CEO contracts are not in equilibrium in our sample (see, e.g., Bushman and Smith 2001). Even if we accept the idea that CEO contracts were not in equilibrium during the 1990s, equity portfolios for CEOs have long been large relative to cash pay, and these portfolios have exhibited cross-sectional variation inconsistent with standard predictions (Demsetz and Lehn 1985).

A third conjecture is that the assumptions underlying the standard agency model do not describe CEO incentives in real institutional settings. For example, Hemmer (2002) discusses how subtle changes in seemingly innocuous assumptions (e.g., the assumption that performance measure variance is not informative about the agent's action) may reverse predictions commonly associated with the agency literature. While Hemmer's (2002) analysis is restricted to a single performance measure, we conjecture that his insights extend naturally to multiple performance measures. For example, if the sensitivity of price to CEO actions increases sufficiently with variance, then a standard agency model might predict a positive association between the relative weight on stock return and *var(ret)* in total compensation. Under this last explanation, our results reject the common assumption that relative performance measure variance proxies for relative performance measure noise. If this explanation is correct, then models characterized by the agent taking actions that affect the mean of the distribution, but not the variance, may fail to describe the relative weights on performance measures used to provide CEO incentives (Hemmer 2002).

We also question whether the assumptions underlying the standard agency model sufficiently incorporate project selection on the part of the CEO. For example, some studies argue that the CEO has much better information about the nature and characteristics of the investment opportunity set than do shareholders and the board of directors. Thus, the primary agency problem is to motivate optimal project selection (e.g., Demsetz and Lehn 1985; Smith and Watts 1992; Prendergast 2002). This implies that firms may use stronger incentives when shareholders are more uncertain about the CEO's choice of projects. Previous studies, such as Demsetz and Lehn (1985) and Core and Guay (1999), interpret a positive association between CEO equity incentives and return volatility, such as that shown in Columns (4) and (5) of Table 4, as support for this hypothesis. This interpretation assumes that return variance is a proxy for uncertainty about the CEO's choice of projects. The problem with reconciling this argument with our analysis, however, is that it only offers an explanation for variation in price-based incentives, not the *relative* weight placed on multiple performance measures. Consequently, like the argument about relative variances above, this argument offers no explanation why cash pay incentives vary in an opposite fashion to total compensation incentives.

VI. CONCLUSION

This paper highlights an issue in the accounting literature that has been generally overlooked since Antle and Smith (1986) and Baker (1987). A typical CEO holds a large portfolio of stock and options. Compared to incentives provided by annual total pay, this portfolio provides a CEO with the majority of his or her total incentives. The significance of this fact is that theory predicts that incentives and rewards for CEOs are based on the relative weight on performance measures in CEOs' total compensation, not the weight on a single component of compensation, such as cash pay. The major contribution of our work is that, although we confirm prior evidence that CEO cash pay varies consistent with standard predictions in the literature, we find that the remaining components of CEO compensation (other annual pay and stock and option holdings) do not.

In the current literature there are arguments and analyses that reconcile our results on total compensation. An argument that reconciles why cash pay varies in a manner that is different from total compensation, however, is more elusive. One explanation is that bonus plans are actually designed to provide incentives for lower-level executives, not the CEO. Consistent with this explanation, the median executive bonus plan in Murphy (2000) includes 122 employees in addition to the CEO. Contracting frictions in real institutional settings might make it efficient to have the CEO's bonus plan parallel that of lower-level executives. For example, it may be more efficient to have a bonus plan for the CEO that parallels that of lower-level employees than it is to explain to employees and investors that the CEO's annual pay differs from that of the rest of the organization because she receives significant incentives from her equity portfolio. If this were true, then the reason that CEO cash compensation varies consistent with standard agency predictions is that the variation in the CEO's cash bonus actually reflects optimal contracting with executives lower in the hierarchy.

The provocative idea suggested by this observation is a previously undocumented benefit to parallel contracting arrangements in hierarchical organizations. The organization as a whole can be more efficient with a contracting device, such as a bonus plan, that motivates some employees in the hierarchy, despite the fact that it is superfluous for others (e.g., the CEO). What may be required is a comprehensive examination of contract efficiency across all elements of the hierarchy, and not simply one member. We believe that this issue is especially promising for future research on performance measurement.

REFERENCES

Abowd, J., and D. Kaplan. 1999. Executive compensation: Six questions that need answering. *The Journal of Economic Perspectives* 13 (4): 145–168.

Aggarwal, R., and A. Samwick. 1999. The other side of the trade-off: The impact of risk on executive compensation. *The Journal of Political Economy* 107 (1): 65–105.

Antle, R., and A. Smith. 1986. An empirical investigation of the relative performance evaluation of corporate executives. *Journal of Accounting Research* 24 (1): 1–39.

Baber, W., S. Janakiraman, and S. Kang. 1996. Investment opportunities and the structure of executive compensation. *Journal of Accounting & Economics* 21 (3): 297–318.

———, S. Kang, and K. Kumar. 1998. Accounting earnings and executive compensation: The role of earnings persistence. *Journal of Accounting & Economics* 25 (2): 169–193.

Baker, G. 1987. Discussion of an analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research* 25 (Supplement): 126–129.

———, and B. Hall. 2002. CEO incentives and firm size. *Journal of Labor Economics* (forthcoming).

Banker, R., and S. Datar. 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* 27 (1): 21–39.

Benston, G. 1985. The self-serving management hypothesis: Some evidence. *Journal of Accounting & Economics* 7 (1–3): 67–83.

Bettis, J., J. Bizjak, and M. Lemmon. 2001. Managerial ownership, incentive contracting, and the use of zero-cost collars and equity swaps by corporate insiders. *Journal of Financial and Quantitative Analysis* 36 (3): 345–370.

Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *The Journal of Political Economy* 81 (3): 637–654.

Boschen, J., and K. Smith. 1995. You can pay me now and you can pay me later: The dynamic response of executive compensation to firm performance. *The Journal of Business* 68 (4): 577–608.

———, A. Duru, L. Gordon, and K. Smith. 2003. Accounting and stock price performance in dynamic CEO compensation arrangements. *The Accounting Review* 78 (1): 143–168.

Bushman, R., and R. Indjejikian. 1993. Accounting income, stock price, and managerial compensation. *Journal of Accounting & Economics* 16 (1-3): 3–23.

———, ———, and A. Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting & Economics* 21 (2): 161–193.

———, Q. Chen, E. Engel, and A. Smith. 2000. The sensitivity of corporate governance systems to the timeliness of accounting earnings. Working paper, University of Chicago, Chicago, IL.

———, and A. Smith. 2001. Financial accounting information and corporate governance. *Journal of Accounting & Economics* 32 (1-3): 237–333.

Core, J., and W. Guay. 1999. The use of equity grants to manage optimal equity incentive levels. *Journal of Accounting & Economics* 28 (2): 151–184.

———, and ———. 2001. When contracts require risk-averse executives to hold equity: Implications for option valuation and relative performance evaluation. Working paper, University of Pennsylvania, Philadelphia, PA.

———, and ———. 2002a. Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting Research* 40 (3): 613–630.

———, and ———. 2002b. The other side of the trade-off: The impact of risk on executive compensation: A revised comment. Working paper, University of Pennsylvania, Philadelphia, PA.

Demsetz, H., and K. Lehn. 1985. The structure of corporate ownership: Causes and consequences. *Journal of Political Economy* 93 (6): 1155–1177.

Feltham, G., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal and agent relations. *The Accounting Review* 69 (3): 429–453.

Garen, J. 1994. Executive compensation and principal-agent theory. *The Journal of Political Economy* 102 (6): 1175–1199.

Gaver, J., and K. Gaver. 1993. Additional evidence on the association between the investment opportunity set and corporate financing, dividend, and compensation policies. *Journal of Accounting and Economics* 16 (1-3): 125–160.

Hall, B., and J. Liebman. 1998. Are CEOs really paid like bureaucrats? *Quarterly Journal of Economics* 113 (3): 653–691.

———, and K. Murphy. 2002. Stock options for undiversified executives. *Journal of Accounting and Economics* 33 (1): 3–42.

Hemmer, T., S. Matsunaga, and T. Shevlin. 1996. The influence of risk diversification on the early exercise of employee stock options by executive officers. *Journal of Accounting and Economics* 21 (1): 45–68.

———. 2002. On the not so obvious relation between risk and incentives and principal-agent relations. Working paper, University of Chicago, Chicago, IL.

Holmstrom, B. 1979. Moral hazard and observability. *The Bell Journal of Economics* 10 (1): 74–91.

———. 1982. Moral hazard in teams. *The Bell Journal of Economics* 13 (2): 324–340.

———, and P. Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55 (2): 303–328.

Huddart, S., and M. Lang. 1996. Employee stock option exercises: An empirical analysis. *Journal of Accounting and Economics* 21 (1): 5–43.

Ittner, C., D. Larcker, and M. Rajan. 1997. The choice of performance measures in annual bonus contracts. *The Accounting Review* 72 (2): 231–255.

Janakiraman, S., R. Lambert, and D. Larcker. 1992. An empirical investigation of the relative performance evaluation hypothesis. *Journal of Accounting Research* 30 (1): 53–69.

Jensen, M., and K. Murphy. 1990. Performance pay and top-management Incentives. *The Journal of Political Economy* 98 (2): 225–264.

Jin, L. 2002. CEO compensation, diversification, and incentives. *Journal of Financial Economics* 66 (1): 29–63.

Lambert, R., and D. Larcker. 1987. An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research* 25 (Supplement): 85–125.

———, ———, and R. Verrecchia. 1991. Portfolio considerations in valuing executive compensation. *Journal of Accounting Research* 29 (1): 129–149.

———. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32 (1-3): 3–87.

Mehran, H. 1995. Executive compensation structure, ownership, and firm performance. *Journal of Financial Economics* 38 (2): 163–184.

Milbourn, T. 2003. CEO reputation and stock-based compensation. *Journal of Financial Economics* 68 (2): 233–262.

Murphy, K. 1985. Corporate performance and managerial remuneration: An empirical analysis. *Journal of Accounting and Economics* 7 (1-3): 11–42.

———. 1999. Executive compensation. In *Handbook of Labor Economics,* edited by O. Ashenfelter, and D. Card, D., Vol. 3. Amsterdam, The Netherlands: North-Holland.

———. 2000. Performance standards in incentive contracts. *Journal of Accounting and Economics* 30 (3): 245–278.

Paul, J. 1992. On the efficiency of stock-based compensation. *The Review of Financial Studies* 5 (3): 471–502.

Prendergast, C. 2002. The tenuous tradeoff between risk and incentives. *Journal of Political Economy* 110 (5): 1071–1102.

Rogers, W. 1993. Regression standard errors in clustered samples. In *Stata Technical Bulletin Reprints,* Vol. 3., 83–94. College Station, TX: Stata Press.

Sloan, R. 1993. Accounting earnings and top executive compensation. *Journal of Accounting and Economics* 16 (1-3): 55–100.

Smith, C., and R. Watts. 1992. The investment opportunity set and corporate financing, dividends, and compensation policies. *Journal of Financial Economics* 32 (3): 263–292.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (4): 817–838.

II. AUDITOR RISK ASSESSMENTS

SAS No. 47 (AICPA 1983a) indicates that auditors should assess inherent risk and control risk at the account balance or class-of-transactions level for purposes of determining the nature, timing, and extent of substantive tests. The risk model is given by the following equation (AU §350.48):[1]

$$AR = IR \times CR \times AP \times TD$$

where:

AR = audit risk;
IR = inherent risk;
CR = control risk;
AP = analytical procedures risk; and
TD = detection risk from tests of details.

The auditor first determines the desired level of audit risk, and then assesses inherent risk, control risk, and analytical procedures risk to determine detection risk and the corresponding extent of substantive testing to be performed.[2] In planning the level of substantive tests, auditors focus on inherent risk and control risk because the combination of $IR \times CR$ equals the likelihood of error in the accounts, prior to the auditor's testing.[3]

Risk assessments can be made in either numerical or nonnumerical terms. Consistent with Houston et al. (1999), all three firms in our study used nonnumerical assessments of risk. The firms in our study differ in the number of levels used in assessing risk, and there were some changes in risk assessment practices across the two time periods we examined.

Risk Assessment Levels

Risk is commonly assessed using two or three measurement levels. Two-level assessments generally take the form of rely/not rely for control risk, and risk/no risk for inherent risk assessments. The AICPA (1999) *Audit Sampling* Audit Practice Release provides four levels for the combined assessment of inherent risk and control risk: (1) maximum, (2) slightly below maximum, (3) moderate, or (4) low.

Only one of the three firms in our sample (Firm C) used a low-medium-high evaluation scheme for both inherent risk and control risk. One of the firms (Firm A) used only two levels for both control and inherent risk. The other firm (Firm B) used three levels for inherent risk, and changed from two assessment levels for control risk in 1994 to three levels in 1999.[4] Our primary focus is on changes in risk assessment levels within firms across time, rather than a comparison across firms. We did not see evidence that firms provided guidance to auditors as to the ranges of probability associated with linguistic assessments. *Audit Sampling* (AICPA 1999) also does not provide probability estimates associated with each assessment level.

---

[1] SAS No. 47 discusses audit risk as consisting of three components: inherent risk, control risk, and detection risk. The formula in the appendix to SAS No. 39 (AICPA 1981) includes an additional component for analytical procedures risk.

[2] There are additional risks that can impact inherent risk and control risk. For example, SAS No. 82 (AICPA 1997) and its successor, SAS No. 99 (AICPA 2002), require auditors to assess the potential for fraud. Bedard and Johnstone (2002) examine whether earnings management and corporate governance risk are associated with increased audit hours and fees. Many large audit firms also consider the effect of client business risk (Bell et al. 1997).

[3] In examining the relationship between inherent and control risk assessments and sample sizes, we do not control for the results of analytical procedures or overall audit risk. We did not see any evidence that the firms in our study made an assessment of analytical procedures risk similar to the assessment of inherent risk and control risk, or directly included the results of analytical procedures in their determination of sample sizes. We also do not have information on overall audit risk assessments.

Limited evidence exists regarding risk assessment levels currently used in practice. Waller (1993) found that auditors at one firm tended to assess control risk at the maximum level, and Mock and Wright (1993) and O'Keefe et al. (1994) also found low reliance on controls. Houghton and Fogarty (1991) emphasize the importance of inherent risk. The PAE also indicates that firms are increasing the emphasis on inherent risk in determining the nature and extent of substantive tests. Increased emphasis on risk assessments, as well as improvements in client controls, could result in lower risk assessments than documented in previous studies, reflecting evolution in the use of the audit risk model.

In addition, competitive factors can influence auditors' risk assessments. PAE focus group participants suggested that time pressures place stress on the audit process and have the potential to adversely affect audit quality (PAE 2000, 106). Kachelmeier and Messier (1990) perform an experiment using the nonstatistical sampling formula in the AICPA (1983b) *Audit Sampling* guide. They detect a tendency among auditors to "work backward" to achieve sample sizes smaller than those suggested by the guide. As a result, auditors may understate the level of risk to achieve a desired sample size. Messier et al. (2001) find that this working-backward phenomenon persists when auditors are asked to determine sample sizes using the revised *Audit Sampling* guide (AICPA 1999), although sample sizes are smaller using the revised guide and closer to auditors' intuitive judgments. In response to fee competition, Houston (1999) finds that auditors reduce audit hours and risk assessment levels.

As the result of changes in risk assessment practices, controls, and increased competition, we expect that risk assessment levels will decrease over time. Our first research question addresses changes in risk assessment levels between the two sample periods.

RQ1: Have auditors' risk assessment levels changed over time?

We consider alternative explanations for such changes in our discussion of results.

Sample Sizes

We first determine each firm's methodology for determining sample sizes. All three firms used nonstatistical methods to determine sample sizes, similar to the nonstatistical sample size formula in the AICPA (1983, 1999) *Audit Sampling* guides:

$$\frac{\textit{Population's recorded amount}}{\textit{Tolerable misstatement}} \times \textit{Assurance factor} = \textit{Sample size.}$$

The assurance factor in the 1999 guide is based explicitly on combined inherent and control risk. In the 1983 guide, the assurance factor is based on the expectation of error in the account, and the desired degree of assurance from the test, factors that at least partially correspond to inherent risk and control risk.

---

[4] The three-level assessment of control risk for this firm was not a low-medium-high assessment. The third level fell between a rely/not rely classification and was used to indicate a reduced level of reliance on controls consistent with reduced testing of controls.

The firms make a few modifications in applying this sampling formula. First, *Tolerable misstatement* is replaced with *Planning materiality* in the calculation of sample sizes. The formula can also be modified so that the *Assurance factor* appears in the denominator. As a result, the formula as applied by these firms is similar to the following:

$$\frac{\textit{Population's recorded amount}}{(\textit{Planning materiality/Assurance factor})} = \textit{Sample size.}$$

In general, the assurance factors used by the firms are smaller than those in the 1999 sampling guide. The combination of lower assurance factors, and the use of planning materiality, rather than tolerable misstatement, results in sample sizes that are smaller than those suggested by the 1999 *Audit Sampling* guide.

Anecdotal evidence suggests that sample sizes have been decreasing. For example, in an archival study of error projection, Elder and Allen (1998) suggest that sample sizes are comparatively small, and the PAE (2000) identified at least some cases involving inadequate sample sizes. The second research question addresses whether sample sizes have in fact decreased between two distinct time periods (1994 and 1999), controlling for the same firms and composition of clients.

RQ2: Have sample sizes decreased over time?

A number of environmental changes could result in lower sample sizes in the later time period. These changes include increased reliance on information systems, increased reliance on and testing of controls, changing importance of sampled accounts like receivables and inventories, and responses to a competitive environment. We control for risk assessment levels to determine whether sample sizes have decreased incrementally to any reduction associated with changes in risk-assessment levels. We consider alternative explanations in our discussion of results.

Relationship between Risk Assessments and Sample Sizes

Previous studies (e.g., O'Keefe et al. 1994; Hackenbrack and Knechel 1997; Johnstone and Bedard 2001; Bell et al. 2001) find a relation between risk assessment and audit hours. These studies often use overall or preplanning risk measures rather than actual inherent and control risk assessments, and do not directly examine whether increased risk is associated with increased sample sizes. Little evidence exists as to firm practices in applying the audit risk model (Dusenbury et al. 1996), and whether risk assessments are related to auditors' evidence decisions. The PAE (2000, 37) recommended that the Auditing Standards Board develop more definitive guidance on linking substantive tests to risk assessments, including reducing the incidence of inadequate sample sizes and variations in sample sizes in similar circumstances.

Mock and Wright (1993, 1999) examine the relationship between auditor risk assessments and the nature and extent of audit procedures using judgments extracted from auditors' workpapers. They find that risk assessments affect planned audit hours, but not the nature of the audit procedures. Bedard (1989) also finds little variation in the number of audit procedures. When audit plans are adjusted, improvement in controls and past favorable results are most often cited as reasons. In a behavioral study of partners and managers, Houston et al. (1999) find that changes in audit fees and investment are sensitive to perceived changes in the risk of errors and fraud.

The risk model indicates that the amount of audit evidence should be related to risk assessments. However, the concerns raised by the PAE (2000) suggest that there may be an insufficient linkage between risk assessments and sample sizes. These concerns are corroborated in a verbal protocol analysis study conducted by Wright and Bedard (2000) in which they find no relationship between inherent risk factors and the extent of audit tests. We are unaware of studies that directly address whether sample sizes or other direct measures of the quantity of evidence are related to risk assessments. The third research question addresses the relationship between risk assessments and sample sizes:

RQ3: Is there a relationship between auditor risk assessments and sample sizes?

Other research has emphasized the importance of inherent risk in identifying errors (Houghton and Fogarty 1991). O'Keefe et al. (1994) find a significant relation between audit hours and inherent risk, but not control risk. Houston et al. (1999) find that auditors place more emphasis on control risk. We address whether the audit firms in our study place comparatively greater emphasis on inherent risk or control risk in their sample size decisions. We first assess the weight given each risk model component in the firm's sampling methodology. We then empirically examine the importance of each component in a regression of sample size on risk model components.

Additional Factors Influencing Sample Size Decisions

In addition to risk assessment factors, sample sizes in tests of balances likely depend upon the dollar amount of the account balance, materiality, type of test, and time period. We expect sample sizes to be positively related to the account balance, and negatively related to materiality. We measure materiality based on planning materiality. We use planning materiality, rather than tolerable misstatement, because some of the firms in our study base tolerable misstatement on the auditor's risk assessments.

We also add an indicator variable for inventory test counts. Audit firms tend to select larger sample sizes for test counts, because of the difficulty of going back and reperforming additional tests. Because these tests involve larger sample sizes, they may be more sensitive to risk. However, because it is difficult to specify the dollar amount of the population and amounts tested, inventory test counts are more frequently determined judgmentally, rather than by formula, which may reduce the relation between risk and sample sizes. Because sample size decisions depend partially on the audit test, we also run separate sample size regression models by type of test, including inventory test counts and accounts receivable confirmation tests.

III. RESEARCH DESIGN AND DATA DESCRIPTION

Data for this study were collected at the offices of two Big 5 audit firms and one large regional audit firm. Each office was located in a different geographic region of the U.S. After reviewing data from sampling applications in the audit files, we gained an understanding of each firm's methodology for determining sample sizes, and the effect of risk assessments on sample sizes. This information was supplemented by discussions with firm personnel, and in some cases, reviews of firm audit manuals.

Sample size and risk assessment information was collected for 235 sampling applications from the accounts receivable and inventory workpapers for a sample of 65 audits for fiscal year 1993 or 1994 (hereafter the 1994 sample). We collected an additional 197 sampling applications for 53 audits in fiscal year 1999 or 2000 approximately five years later (hereafter the 1999 sample). We attempted to obtain data in the later period for the same clients that were examined in the earlier period. However, many clients from the 1994 sample were no longer clients when the 1999 sample was collected or the working papers were not available in the office during the time of our visits. The 1999 sample includes 35 clients common to the earlier period, and an additional 18 clients that had not been tested in the earlier period. We report results for both the full set of observations, and the reduced sample of 35 clients common to both periods.

The common sample of 35 clients maximizes comparability for tests of changes in risk assessments and sample sizes over time. However, client changes to address risk and auditor familiarity with the client likely result in reduced risk and sample sizes over time for any given client. We therefore report results for the full sample of 53 clients in addition to the common sample of 35 clients to provide evidence on a broader sample of clients that mitigates the effects of auditor familiarity and changes in client characteristics over time. We also perform additional analysis to address whether changes in client characteristics explain reductions in risk and sample sizes.

We collected data from audit files at each audit firm's office. We agreed not to disclose the names of the firms or their clients, but otherwise were not restricted in our data collection procedures. Firm representatives selected audit working papers on the basis of their availability in the office at the time of our visit. The selected clients had accounts receivable and/or inventory, and were primarily manufacturers or distributors. Although representatives of the firms selected the clients, there was no indication that there was any attempt to screen clients or otherwise restrict our access to data.

Descriptive information for the 1994 and 1999 samples is included in Panels A and B of Table 1. The average client had $60 million in sales and $42 million in assets in the 1994 sample, as compared to average sales of $144 million and average assets of $84 million in the later period. There is some variation in client size across the three audit firms, and regional Firm B has the smallest clients based on sales and assets.

Risk Assessments

Table 2 reports the inherent and control risk assessments for each audit firm for 1994 and 1999. Inherent risk assessments are presented in Panel A and control risk assessments are presented in Panel B. Our analysis is based on account balance risk assessments. Although the firms in our study assessed risk by assertion within account, in almost all cases, the risk assessments were the same across assertions. This similarity of risk assessments across assertions is consistent with the findings in Waller (1993).[5]

Panel A of Table 2 indicates that inherent risk assessments are significantly lower in 1999 than in 1994 for Firm B, Firm C, and the overall sample. There is no significant difference between periods in Firm A's inherent risk assessments. Firm A uses a two-level classification of specific risk/no risk identified to assess inherent risk, and few risks were identified in either period. Risk assessments reflect auditor judgment. As a result, it is difficult to measure whether a risk assessment is "too high" or "too low." In general, the risk assessments identified in this study are lower than those identified in previous research (Waller 1993).

Of particular interest is the lack of risks in receivables and inventory identified by Firm A using the two-level classification scheme. One possibility is that this firm has unusual

---

[5] For one audit firm, inherent risk for valuation of accounts receivable was frequently assessed higher than inherent risk for other assertions. The higher assessed inherent risk for valuation appeared to be related to concerns over net realizable value. Another firm assessed net realizable value separately from accuracy for the valuation assertion.

TABLE 1
Descriptive Information for Sample Audit Clients

Panel A: 1994 Sample

	Mean (std. deviation) (000s omitted)			
Financial Measure	(n = 26) **Firm A**	(n = 19) **Firm B**	(n = 20) **Firm C**	(n = 65) **Overall**
Sales	$85,492.4 (142,460.5)	$34,151.8 (29,496.4)	$49,142.8 (97,753.9)	$59,760.0 (108,533.4)
Total Assets	42,370.7 (38,662.2)	28,276.4 (42,587.0)	51,224.8 (95,978.1)	41,883.1 (64,633.6)
Net Income	1,443.0 (2,691.2)	2,011.9 (5,328.6)	3,235.6 (6,212.9)	2,213.5 (4,834.0)
Accounts Receivable	8,593.0 (9,482.1)	13,110.3 (25,261.1)	6,169.7 (10,602.2)	8,894.0 (15,458.0)
Inventory	10,492.0 (14,323.5)	12,858.4 (23,521.3)	17,709.1 (53,465.8)	13,558.5 (34,265.0)

Panel B: 1999 Sample

	Mean (std. deviation) (000s omitted)			
Financial Measure	(n = 27) **Firm A**	(n = 14) **Firm B**	(n = 12) **Firm C**	(n = 53) **Overall**
Sales	$201,183.0 (238,869.0)	$33,618.9 (23,654.6)	$126,743.3 (154,628.2)	$143,545.1 (199,980.5)
Total Assets	99,966.0 (95,292.3)	16,442.1 (11,319.4)	119,286.5 (181,816.4)	84,480.5 (114,258.4)
Net Income	6,658.5 (11,027.3)	3,914.6 (9,187.3)	5,353.4 (5,072.4)	5,662.9 (9,516.6)
Accounts Receivable	20,738.5 (21,384.8)	7,546.4 (11,304.4)	16,065.0 (17,430.0)	16,283.9 (18,949.2)
Inventory	28,310.5 (44,340.0)	10,568.8 (25,091.0)	46,709.1 (91,210.1)	28,009.7 (54,867.4)

clients. We use the percentage of tests with errors and the projected errors uncovered in the testing as a measure of the actual risk in the audit test. Both measures were higher for Firm A than for the other two audit firms, suggesting that the difference in risk assessments is not attributable to client differences.

Panel B of Table 2 indicates significantly lower control risk assessments in 1999 than in 1994 for Firm A and Firm C. Firm A uses a two-level rely/not rely classification of control risk, and Firm C uses a three-level low/medium/high classification of control risk. Firm B changed from two levels of control risk to three levels of control risk between 1994 and 1999. It is therefore not possible to conduct a meaningful statistical test for changes

TABLE 2
Risk Assessments by Audit Firm and Sample Period

	Full Sample (Matched Sample) 1994 Inherent Risk Assessment			Full Sample (Matched Sample) 1999 Inherent Risk Assessment			
Firm	**Low**	**Med**	**High**	**Low**	**Med**	**High**	**χ^2-stat.**
Panel A: Inherent Risk Assessments: 1994 and 1999 Samples							
A	103	NA	2	107	NA	1	.37
	(63)	(NA)	(1)	(70)	(NA)	(1)	(.06)
B	3	44	19	19	14	13	25.5***
	(0)	(21)	(12)	(10)	(9)	(10)	(14.8)***
C	34	20	1	23	1	3	12.2***
	(13)	(8)	(0)	(17)	(1)	(3)	(9.0)**
Total	140	64	22	149	15	17	26.7***
	(76)	(29)	(13)	(97)	(10)	(14)	(11.8)***
Panel B: Control Risk Assessments: 1994 and 1999 Samples							
A	82	NA	23	106	NA	2	20.7***
	(53)	(NA)	(11)	(70)	(NA)	(1)	10.4***
B	57	NA	9	17	17	12	NA#
	(28)	(NA)	(5)	(16)	(8)	(5)	(NA)#
C	23	21	10	23	1	3	14.6***
	(11)	(5)	(4)	(17)	(1)	(3)	(4.1)
Total	162	21	42	146	18	17	NA#
	(92)	(5)	(20)	(103)	(9)	(9)	(NA)#

*, **, *** Indicates χ^2-statistic significant at the .10 (.05) (.01) levels, respectively.
χ^2-statistic not applicable because Firm B changed from two-level to three-level control risk assessments between 1994 and 1999.
Full sample includes all observations. Matched sample includes only observations for the 35 clients common to both periods.
Medium is shown as NA when the firm used two levels for risk assessment.
For inherent risk, two-level risk assessment is based on risk/no risk identified.
For control risk, two-level risk assessment is based on rely/not rely on controls.

in control risk assessments for Firm B. The firm used a two-level below maximum/maximum assessment in the 1994 period.[6] The three levels in the 1999 period can be classified as below maximum/slightly below maximum/maximum. Based on these terms and discussion with firm personnel, the pattern does not reflect a significant reduced reliance on controls. The use of the term "slightly below maximum" appears to reflect reduced tests of controls.

In addition to lower assessments of control risk, the audit firms appear to be reducing tests of controls. For example, Firm A changed to rotational testing of controls in the 1999 period. In some cases, the auditors further relied on controls if they were subject to testing during the current-year audit.

---

[6] The two control risk levels used by Firm A and Firm B appear to have the same interpretation, but the difference in terminology is interesting. Firm A uses the term "rely on controls," whereas Firm B indicates that control risk is "below the maximum," which is consistent with reliance on controls. Neither firm expresses the degree of reliance.

Control risk assessments depend on both the existence of controls and the auditor's decision to rely on those controls. The data suggest a high and increasing reliance on controls. This extensive reliance on controls differs from the findings in Waller (1993), Mock and Wright (1993), and O'Keefe et al. (1994). Control risk was set at the maximum for the majority of clients in these studies. Those studies are based on data from individual audit firms, so it is difficult to determine whether auditors have increased their reliance on controls. However, there are several reasons why auditors have likely increased their reliance on controls. In the past, auditors may have assessed control risk as high and eliminated the documentation and tests of controls. However, SAS No. 55 (AICPA 1988) requires auditors to obtain an understanding of internal control. This requirement, along with rotational testing of controls and a likely reduction in sample sizes for tests of controls, have likely made it more cost effective to test and rely on controls.

Sample Size

Panel A of Table 3 reports mean sample sizes by time period and by firm. The average sample size decreased from 30.7 in 1994 to 26.0 in 1999 for the full sample. The decrease is not statistically significant for the full sample, and is marginally significant for the matched sample. However, further scrutiny indicates that this overall apparent stability in sample sizes is influenced heavily by Firm A. Firms B and C exhibit a statistically significant decrease in sample sizes from 1994 to 1999, although the decrease is not significant for the matched sample for Firm B. Across firms, 1994 sample sizes were significantly lower in Firm A than in Firms B and C ($t = 3.63$ $p < .001$ for full sample; $t = 1.88$, $p = .06$ for matched sample). However, in 1999 there is no significant difference in sample sizes between the three firms ($t = .78$, $p = .44$ for full sample; $t = .65$, $p = .52$ for matched sample). The descriptive sample size data therefore suggests that Firms B and C decreased sample sizes from 1994 to 1999, toward the smaller sample sizes that characterized Firm A in both periods.

Multiple observations from the same client are not independent. Accordingly, we performed a supplemental analysis of sample sizes reported in Table 3 by collapsing multiple observations from each audit client into a single average sample size for that client for both the 1994 and 1999 periods. This analysis is somewhat more supportive of a decline in sample sizes than the results reported in Table 3, Panel A. For all firms, the t-statistic for the change in sample size for the full sample becomes marginally significant, and the reduced sample increases in significance to the .05 level. For Firm B and Firm C, the decrease in sample sizes for the full sample is significant at the .01 level. The reduced sample becomes marginally significant at the .10 level for Firm B and remains significant at the .05 level for Firm C.

Panel B of Table 3 reports mean sample sizes by type of test and time period for the overall sample. The average sample size for inventory price tests decreased significantly in 1999 as compared to 1994 for the full sample. One possible explanation for the relatively larger reduction in inventory testing is increased reliance on improved inventory control systems. The average sample size for accounts receivable confirmations and inventory test counts also declined. The decrease in accounts receivable confirmations is marginally significant for the full sample, and the decline in inventory test counts is marginally significant for the matched sample. The average sample size for cutoff tests and other tests increased slightly, although the increase is not statistically significant at conventional levels. As expected, sample sizes were much larger for inventory test counts.

TABLE 3
Sample Sizes by Audit Firm, Audit Test, and Sample Period

Panel A: Mean Sample Size by Time Period and By Firm

	Mean Sample Size [# of obs.]—Full Sample (Mean Sample Size) [# of obs.]—Matched Sample			
Audit Firm	**1994**	**1999**	**Combined**	**t-statistic**
Firm A	21.0 [n = 109]	27.4 [n = 121]	24.4 [n = 230]	−1.39
	(23.1) [n = 68]	(21.9) [n = 78]	(22.5) [n = 146]	(0.32)
Firm B	35.3 [n = 66]	21.8 [n = 46]	29.7 [n = 112]	2.17**
	(31.5) [n = 33]	(24.9) [n = 29]	(28.4) [n = 62]	(0.77)
Firm C	43.1 [n = 60]	26.7 [n = 30]	37.6 [n = 90]	2.01**
	(37.4) [n = 25]	(23.9) [n = 24]	(30.8) [n = 49]	(1.70)**
All Firms	30.7 [n = 235]	26.0 [n = 197]	28.6 [n = 432]	1.25
	(28.2) [n = 126]	(22.9) [n = 131]	(25.5) [n = 257]	(1.53)*

Panel B: Mean Sample Size by Time Period and Type of Test—All Firms

	Mean Sample Size [# of obs.]—Full Sample (Mean Sample Size) [# of obs.]—Matched Sample			
Type of Test	**1994**	**1999**	**Combined**	**t-statistic**
Accounts Receivable Confirmations	22.1 [n = 57]	16.6 [n = 47]	19.6 [n = 104]	1.61*
	(17.5) [n = 31]	(14.1) [n = 29]	(15.9) [n = 60]	(1.08)
Inventory Test Counts	54.9 [n = 62]	43.4 [n = 65]	49.0 [n = 127]	1.04
	(47.5) [n = 34]	(34.0) [n = 43]	(40.0) [n = 77]	(1.50)*
Inventory Price Tests	25.6 [n = 67]	17.8 [n = 49]	22.3 [n = 116]	2.17**
	(27.1) [n = 35]	(19.1) [n = 35]	(23.1) [n = 70]	(1.46)*
Cutoff Tests	15.1 [n = 31]	17.0 [n = 31]	16.0 [n = 62]	−0.61
	(14.8) [n = 17]	(18.1) [n = 20]	(16.6) [n = 37]	(−0.69)
Other Tests	19.8 [n = 18]	23.8 [n = 5]	20.7 [n = 23]	−0.43
	(12.8) [n = 5]	(25.8) [n = 4]	(18.6) [n = 9]	(−1.28)
Overall Mean	30.7 [n = 235]	26.0 [n = 197]	28.6 [n = 432]	1.25
	(28.2) [n = 126]	(22.9) [n = 131]	(25.5) [n = 257]	(1.53)*

Panel C: Mean Sample Size by Time Period and Type of Test—Firms B and C

	Mean Sample Size [# of obs.]—Full Sample (Mean Sample Size) [# of obs.]—Matched Sample			
Type of Test	**1994**	**1999**	**Combined**	**t-statistic**
Accounts Receivable Confirmations	27.4 [n = 31]	15.8 [n = 19]	23.0 [n = 50]	2.41**
	(20.5) [n = 13]	(12.0) [n = 10]	(16.8) [n = 23]	(2.06)**
Inventory Test Counts	72.6 [n = 33]	34.2 [n = 24]	56.4 [n = 57]	2.47***
	(58.1) [n = 14]	(34.4) [n = 18]	(44.8) [n = 32]	(1.39)*
Inventory Price Tests	40.0 [n = 42]	28.5 [n = 22]	36.0 [n = 64]	1.74**
	(35.8) [n = 20]	(24.5) [n = 17]	(30.6) [n = 37]	(1.42)*
Cutoff Tests	16.0 [n = 14]	10.1 [n = 7]	14.0 [n = 21]	1.53*
	(17.4) [n = 7]	(9.0) [n = 4]	(14.3) [n = 11]	−(1.32)
Other Tests	19.3 [n = 6]	25.8 [n = 4]	21.9 [n = 10]	−0.53
	(14.3) [n = 4]	(25.8) [n = 4]	(20.1) [n = 8]	(−0.90)
Overall Mean	39.0 [n = 126]	23.7 [n = 76]	33.2 [n = 202]	3.00***
	(34.1) [n = 58]	(24.1) [n = 53]	(29.5) [n = 111]	(1.65)*

---

*, **, *** Significant at the .10 (.05) (.01) levels, respectively, based on a one-tailed test.
Full sample includes all observations. Matched sample includes only observations for the 35 clients common to both periods.

The univariate changes in sample size by type of test reported in Panel B include Firm A. However, as reported above for overall sample sizes, Firm A did not decrease sample sizes in the later period, likely reflecting the fact that Firm A's sample sizes were significantly lower than the other two firms in the earlier period. Accordingly, Panel C of Table 3 includes changes in sample size by type of test for Firms B and C only. For these two firms, sample sizes for accounts receivable confirmations decreased significantly for both the full and matched sample. Sample size decreases for inventory test counts and inventory price tests are significant for the full sample, and marginally significant for the matched sample.

The univariate tests indicate that both Firms B and C decreased sample sizes between the two periods, and that these decreases occurred across each type of test. However, sample sizes are also affected by risk and other factors. The multivariate sample size models presented in the next section control for these additional factors.

IV. MULTIVARIATE TESTS OF THE RELATION BETWEEN SAMPLE SIZES AND RISK

The multivariate sample size model tests for decreases in sample sizes between the two periods after controlling for inherent risk assessments, control risk assessments, population size, materiality, and inventory test counts. We converted the risk assessments to a numerical scale in which 1 = low risk, 2 = medium risk, and 3 = high risk. For firms that specified only two levels of risk, these assessments were coded as 1 = low and 3 = high. The intent of this coding scheme was to maintain consistency with the three-level scheme, but it may tend to understate (overstate) the low (high) risk in the two-level classification. Accordingly, we performed alternate tests using different ordinal rankings, obtaining similar results.[7] We also computed a combined ranking as the product of the two risk measures. This combined measure is consistent with the mathematical structure of the risk model, in which the risk of error is the product of inherent risk and control risk. The combined risk variable is positive and significant, and the model has similar explanatory power to the model with separate assessments of inherent risk and control risk.

Our sample includes multiple observations from the same audit clients. As a result, the error terms are likely to be correlated. To address this concern, we recompute the standard errors and t-statistics after controlling for correlations in error terms using the procedure developed by Froot (1989).[8] Prior accounting studies (e.g., Bamber et al. 1997; Bhattacharya

---

[7] Alternate classification measures capture the likelihood that high risk in a three-level classification scheme suggests greater risk than high risk in a two-level classification scheme. The following alternate classification schemes were tested:

Two-Level	Three-Level	Two-Level	Three-Level
2 = low 4 = high	1 = low 3 = medium 5 = high	1= low 4 = high	1 = low 3 = medium 5 = high

[8] We assume that multiple observations from the same client are not independent for a given time period. However, observations between different clients are assumed to be independent. We run the procedure recommended by Froot (1989) by assuming that multiple observations from the same client in the same period are correlated, while observations from the same client in different time periods and observations from different clients are not correlated. Our indicator variable (1999 year) already captures the time-varying effects of sample size. The significance of most variables is unchanged using the Froot standard errors compared to the OLS estimates. Control risk for the full sample and 1994 period is significant at the .05 and .01 level, respectively, using the OLS standard errors, but is only marginally significant using the Froot standard errors. In the analysis by type of test, the materiality control variable is significant for inventory test counts and price tests using the Froot standard errors, but is only marginally significant using the OLS standard errors.

2001) have used Froot's technique to control for similar issues involving correlated observations.

Combined Firm Analysis

Panel A of Table 4 reports regression results for the combined sample periods. The first column is based on the full sample, but the sample size is 335 due to missing data for other explanatory variables. The second column reports the model for the matched sample of 189 observations for the 35 companies included in both sample periods. The last two columns report the results for the separate 1994 and 1999 samples.

The year indicator variable is negative and significant in both the full and reduced sample, indicating that sample sizes were lower in 1999 than in 1994. Analysis of decreases in sample size by type of test for the full and matched sample are reported in Panels B and C of Table 4, respectively. Individual regressions for each audit firm are reported in Panel A of Table 5.

For the combined period models, the sample size coefficent for inherent risk is positive and significant in both the full and matched sample, while control risk is positive and marginally significant only for the full sample. The coefficient on inherent risk is larger than the coefficient for control risk and the difference is statistically significant. This result is consistent with O'Keefe et al. (1994), who find that hours are significantly related to an inherent risk measure but not control risk, and Houghton and Fogarty (1991), who find that errors are more strongly correlated with inherent risk than control risk. The risk variables are also significant in the separate 1999 and 1994 samples, with the exception of control risk, which is not significant in the 1999 period.

The additional control variables in the model are significant in the expected directions. Sample sizes increase with the size of the sample population, and decrease with the measure of materiality. Sample sizes are significantly larger for inventory tests counts, possibly because of the risks associated with inventory, as well as the difficulty of reperforming these tests at a later date.[9]

Examination of residuals and influence diagnostics identified 13 potentially influential observations. Most of these involved unusually large sample sizes, often for inventory test counts, although a few unusually low sample sizes were also observed. The significance of the results is unchanged when these observations are deleted from the analysis. Because inventory tests counts had larger average sample sizes, and inventory observations are a unique form of audit test, we also reperformed the analysis excluding inventory test counts. When inventory test counts are excluded, the control risk variable increases in significance, and the significance of the remaining variables is unchanged.

We repeat the analysis of sample size by each type of audit test for the full sample and matched sample in Panels B and C, Table 4.[10] These results are generally consistent with the combined tests in Panel A. The year indicator variable is significant for accounts receivable confirmations and inventory price tests for both the full and reduced samples.

Analysis by Audit Firm

Table 5, Panel A reports regressions for each audit firm across both periods of study for the full and matched samples. For Firm A, control risk is significant for both samples.

---

[9] Another explanation for increased sample sizes for inventory test counts is that once the cost investment has been made to visit the client location, performing additional test counts may not be very costly for many types of inventories.

[10] We exclude cutoff and other tests from this analysis because of the limited number of observations for these tests.

TABLE 4
Sample Size Regression Model

Panel A: Combined Periods and Individual 1999 and 1994 Samples

Variable	Pred. Sign	Coefficient (t-statistic) (n = 335) Full Sample Combined Periods	(n = 189) Matched Sample Combined Periods	(n = 196) 1994 Sample	(n = 139) 1999 Sample
Intercept		−7.73	1.00	−21.60	0.51
		(−0.79)	(0.15)	(−1.35)*	(0.11)
Inherent risk	+	15.58	9.68	20.14	7.19
		(3.29)***	(2.76)***	(2.72)***	(3.26)***
Control risk	+	4.43	2.76	8.19	0.85
		(1.30)*	(1.12)	(1.59)*	(0.38)
Population size	+	0.0009	0.0009	0.0016	0.0001
		(5.08)***	(2.65)***	(3.52)***	(5.29)***
Materiality	−	−0.008	−0.007	−0.015	−0.005
		(−3.77)***	(−3.29)***	(−3.68)***	(−3.79)***
Inventory Counts	+	28.64	22.71	30.62	25.03
		(4.76)***	(3.86)***	(3.43)***	(4.48)***
1999 Year	−	−9.33	−6.52	—	—
		(−2.83)***	(−1.97)**		
Adjusted R^2		0.274	0.292	0.292	0.311
F-value		21.97***	13.93***	17.09***	13.47***

Panel B: Sample Size Model by Type of Test—Full Sample

Variable	Pred. Sign	Coefficient (t-statistic) (n = 98) AR Confirmations	(n = 97) Inventory Test Counts	(n = 104) Inventory Price Tests
Intercept		7.14	−11.30	3.21
		(1.25)	(−0.43)	(0.57)
Inherent risk	+	5.15	33.60	7.06
		(1.78)**	(2.87)***	(2.21)**
Control risk	+	3.60	3.79	7.70
		(1.69)*	(0.35)	(2.61)***
Population size	+	0.0009	0.0015	0.0004
		(5.42)***	(4.47)***	(2.69)***
Materiality	−	−0.008	−0.013	−0.005
		(−3.50)***	(−2.50)***	(−3.29)***
1999 Year	−	−9.23	−11.14	−8.57
		(−3.94)***	(−1.24)	(−2.64)***
Adjusted R^2		0.373	0.267	0.194
F-value		12.53***	8.01***	5.96***

(continued on next page)

TABLE 4 (continued)

Panel C: Sample Size Model by Type of Test—Matched Sample

		Coefficient (t-statistic)		
Variable	**Pred. Sign**	**(n = 54) AR Confirmations**	**(n = 56) Inventory Test Counts**	**(n = 60) Inventory Price Tests**
Intercept		13.31	8.83	5.14
		(2.32)**	(0.77)	(0.72)
Inherent risk	+	0.70	22.62	4.32
		(0.25)	(2.51)***	(1.10)
Control risk	+	1.27	−2.96	11.11
		(0.63)	(−0.46)	(2.18)**
Population size	+	0.0007	0.0019	0.0004
		(4.07)***	(2.05)**	(1.73)**
Materiality	−	−0.006	−0.013	−0.005
		(−3.55)***	(−2.81)***	(−2.45)***
1999 Year	−	−7.80	−5.60	−9.38
		(−3.22)***	(−0.70)	(−1.93)**
Adjusted R^2		0.364	0.388	0.175
F-value		7.07***	7.96***	3.51***

*, **, *** Significant at the .10 (.05) (.01) levels, respectively, based on a one-tailed test; t-statistics are calculated using Froot (1989) standard errors.

Full sample includes all observations. Matched sample includes only observations for the 35 clients common to both periods.

Inherent risk is significant for the reduced sample and is marginally significant for the full sample.[11] Firm A exhibited limited variation in risk assessments, and had smaller sample sizes than the other two firms in the 1994 period. Nevertheless, Firm A's risk methodology provides an explicit link between risk and sample size decisions, and the audits with higher assessed risks had larger sample sizes.

The model for Firm B indicates no significant relationship between sample sizes and either inherent risk or control risk. There was no indication in this firm's methodology and audit documentation that sample size decisions were explicitly linked to risk assessments. The indicator variable for year indicates that this firm reduced its sample sizes in 1999, relative to 1994.

The relationship between inherent risk and sample sizes is highly significant for Firm C, but only for the full sample. The coefficient for control risk is positive, but not significant. This firm had the largest average sample size of the three firms in the study, and the indicator variable for year indicates that this firm significantly reduced sample sizes between the two periods.[12]

---

[11] Control risk for Firm A exerts a marginally significant influence on sample sizes in the full sample and is insignificant for the matched sample using OLS standard errors. In the 1999 sample, inherent risk and control risk for Firm A are significant using the Froot standard errors, and insignificant using the OLS standard errors. Use of the Froot standard errors did not impact the significance of the results for Firms B and C.

[12] We cannot completely explain the negative adjusted R^2 for the matched sample for Firm C. One factor affecting adjusted R^2 is the small sample size. The results are also affected by two outlier observations for one client. When these observations are deleted, the adjusted R^2 becomes positive, but is still much smaller than the adjusted R^2 for the full sample and for Firms A and B.

TABLE 5
Sample Size Regression Model by Audit Firm

Panel A: Combined 1994 and 1999 Samples

		Coefficient (t-statistic)					
		Firm A		Firm B		Firm C	
Variable	**Pred. Sign**	**(n = 176) Full Sample**	**(n = 109) Matched Sample**	**(n = 89) Full Sample**	**(n = 48) Matched Sample**	**(n = 70) Full Sample**	**(n = 32) Matched Sample**
Intercept		−2.285	−9.06	22.24	25.42	−28.34	26.33
		(−0.295)	(−1.51)*	(1.81)**	(1.77)**	(−0.81)	(2.24)**
Inherent risk	+	9.38	17.38	1.59	−2.72	38.06	1.02
		(1.46)*	4.09)***	(0.42)	(−0.48)	(1.94)**	(0.10)
Control risk	+	3.69	3.56	−2.42	−0.85	8.05	13.31
		(2.50)***	(2.15)**	(−0.83)	(−0.32)	(1.03)	(1.29)
Population size	+	0.0008	0.0007	0.0007	0.0009	0.0023	0.0018
		(7.99)***	(3.99)***	(0.93)	(1.05)	(2.20)**	(0.91)
Materiality	−	−0.007	−0.007	0.009	0.008	−0.045	−0.050
		(−4.91)***	(−4.13)***	(1.03)	(0.83)	(−1.34)*	(−0.94)
Inventory Counts	+	20.27	18.93	42.51	45.42	23.27	2.24
		(4.09)***	(3.36)***	(3.19)***	(3.05)***	(1.72)**	(0.21)
1999 Year	−	0.71	−0.89	−18.87	−20.29	−22.72	−19.38
		(0.22)	(−0.22)	(−2.24)**	(−2.28)**	(−2.02)**	(−2.33)**
Adjusted R^2		0.295	0.293	0.359	0.449	0.256	−0.02
F-value		13.18***	8.46***	9.20***	7.40***	4.95***	0.91

Panel B: 1994 and 1999 Samples

		Coefficient (t-statistic)					
		Firm A		Firm B		Firm C	
Variable	**Pred. Sign**	**(n = 94) 1994**	**(n = 82) 1999**	**(n = 52) 1994**	**(n = 37) 1999**	**(n = 50) 1994**	**(n = 20) 1999**
Intercept		8.15	−22.08	20.39	5.41	−93.04	18.21
		(1.18)	(−11.29)***	(1.68)	(0.74)	(−1.85)**	(2.69)***
Inherent risk	+	2.02	12.06	1.51	3.90	63.19	2.70
		(0.33)	(5.21)***	(0.42)	(0.98)	(2.75)***	(0.36)
Control risk	+	3.13	17.48	−7.93	−0.43	18.96	3.35
		(2.25)**	(5.64)***	(−1.49)*	(−0.18)	(1.84)**	(0.36)
Population size	+	0.0011	0.0006	0.0012	−0.0000	0.0037	0.0014
		(3.93)***	(6.57)***	(1.03)	(−0.00)	(2.17)**	(4.73)***
Materiality	−	−0.010	−0.005	0.017	0.015	−0.041	−0.047
		(−5.25)***	(−4.02)***	(0.61)	(1.64)*	(−0.87)	(−4.24)***
Inventory Counts	+	9.12	31.38	60.58	18.94	28.95	9.38
		(1.95)**	(3.83)***	(3.08)***	(1.56)*	(1.71)**	(1.21)
Adjusted R^2		0.255	0.372	0.432	0.312	0.399	0.438
F-value		7.38***	10.58***	8.74***	4.26***	7.51***	3.96***

---

*, **, *** Significant at the .10 (.05) (.01) levels, respectively, based on a one-tailed test; t-statistics are calculated using Froot (1989) standard errors.

Full sample includes all observations. Matched sample includes only observations for the clients common to both periods.

Panel B of Table 5 reports the within-firm results for the two sample periods, considered separately. For Firm A, control risk is significant in each period, but inherent risk is significant only for the 1999 period. Consistent with the combined period result, neither inherent risk nor control risk is significant for Firm B. Both inherent risk and control risk are significant for Firm C in the 1994 period. However, in the 1999 period, neither inherent risk nor control risk is significant. Risk assessments and sample sizes are lower in the 1999 period. In addition, Firm C used combined inherent risk and control risk assessments to develop an overall risk assessment, which may have further weakened the relation between risk and sample sizes. However, the results for the 1999 period are based on a relatively small sample size (n = 20) and should be interpreted with caution.

We interpret these results as follows. The existence of a relationship between sample sizes and risk depends on sufficient variation in risk assessments and sample sizes. Also, the existence of a relation between risk and sample sizes is less likely when the decision is based on auditor judgment (Firm B) as opposed to structured guidance (Firms A and C). This supports the recommendation from the PAE to strengthen the relation between risk and sample sizes. Second, even after controlling for decreased risk assessments, the two firms with higher sample sizes in the 1994 sample period (i.e., Firms B and C) decreased sample sizes, consistent with a reduction in substantive sampling over time.

Discussion and Additional Analyses

Reduced sample sizes in 1999 relative to 1994 are consistent with the lower assurance factors in the 1999 *Audit Sampling* guide and the experimental results in Messier et al. (2001). Although lower risk assessments and sample sizes may reflect increased competition, these differences could also reflect genuine changes in client risk, audit efficiency, or changes in audit approaches that involve substitution of other audit procedures for substantive tests of details. We address each of these possibilities based on our knowledge of the each firm's audit approach, discussion with firm personnel, and comparison of audit populations and corresponding tests across periods.

Discussions with firm personnel at Firm B and Firm C support the interpretation that reductions in sample sizes at least partially reflect a response to increased competition. Partners at Firm B indicated that sample size reductions were a competitive necessity. In the 1994 period, individuals at Firm C offered on an unsolicited basis that they were aware that sample sizes were lower at Firm A.[13]

To assess whether lower risk assessments might reflect client changes that support lower inherent and control risk, we first calculate the ratio of receivables to sales and inventory to sales for both periods and by firm to determine whether the relative significance of receivables and inventory decreased during the period of study. For the matched sample (i.e., same clients in 1994 and 1999), receivables increased from 12.1 percent of sales to 13.3 percent of sales, while inventory decreased from 17.0 of sales to 14.5 percent of sales.[14] Neither of these changes is statistically significant. We also gathered data on the amount of errors in the sample applications in each period. For the matched sample, the percentage of tests with errors increased from 44 percent in 1994 to 50 percent in 1999, and the average dollar amount of the errors increased from $14,000 to $28,000. It, therefore, does not appear that smaller sample sizes can be justified by lower error rates.

---

[13] Each firm was aware that we were conducting research at more than one firm, but did not know the identity of the other firms.

[14] One observation had an extremely high level of receivables and inventory compared to sales in the 1994 period. When this observation is excluded, receivables increase from 10.3 percent of sales to 13.0 percent, and inventory decreases from 15.0 percent to 14.2 percent of sales.

Changes in the extent of testing between the two periods may be due to changes in the focus on fraud after SAS No. 82 (AICPA 1997), or to follow a client business risk or strategic system auditing approach such as that described in Bell et al. (1997). We did not have access to all audit files, but we gained a sufficient understanding of each firm's audit approach to address whether the decreases in risk and sample sizes were attributable to changes in audit approaches.

To varying degrees the firms in our sample adopted business risk audit approaches. However, these new audit approaches had not been extensively applied by 1999, and we saw no evidence in the workpapers or from our discussions with client personnel that changes in risk assessments and sample sizes were due to client business risk approaches. We also did not see evidence of other fundamental differences in testing approaches in receivables and inventory, such as increased reliance on systems testing. However, it is possible that reductions in sampling are offset to varying degrees by increases in other sources of assurance.

Shelton et al. (2001) found that SAS No. 82 did not have a significant impact on audit approaches to assessing the risk of fraud. Similarly, Johnstone and Bedard (2001) did not find a significant relationship between fraud risk assessments and audit hours. We did not see evidence of fraud testing in inventory or accounts receivable workpapers, and we did not see evidence of fraud risk assessments impacting sampling plans. However, it is possible that fraud risk assessments impacted inherent risk assessments.

V. SUMMARY, LIMITATIONS, AND CONCLUSIONS

This study examined longitudinal changes in auditors' risk assessments and sample size decisions. We find lower risk assessments for both control risk and inherent risk than previous studies. We also find that risk assessments and sample sizes decreased from 1994 to 1999, based on working papers from offices of three large audit firms with 35 clients common to the two periods studied. Field data do not allow us to definitively identify why sample sizes declined over time. However, anecdotal and supplemental evidence suggests that at least some of the decline in sample sizes can be attributed to competitive pressures in the audit industry.

Consistent with the risk model, sample sizes were associated with measures of both inherent risk and control risk. Further analysis indicated that the relationship between sample sizes and risk assessments varied across firms and time periods. The relationship between control risk and inherent risk and sample sizes was not significant for all firms, and was stronger in the earlier period. The relationship between risk and sample sizes appears to depend upon the availability of firm guidance that explicitly links sample sizes to risk, supporting the call by the Panel on Audit Effectiveness (2000, 29) to strengthen the link between risk and sample sizes. Further research using experimental or archival data is necessary to assess how the number of risk assessment levels and the extent of firm guidance affect risk assessments and sample size decisions.

This study is based upon field research data from three firms and offices. Other firms, and other offices within the same firm, may have different sample sizes and risk assessment practices. Additional evidence is necessary to determine whether risk assessments and sample sizes have continued to decline, or if the recommendations of the Panel on Audit Effectiveness and fraud guidance in SAS No. 99 have affected auditors' risk assessments and sample size decisions.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 1981. *Audit Sampling*. Statement on Auditing Standards No. 39. New York, NY: AICPA

———. 1983a. *Audit Risk and Materiality in Conducting an Audit*. Statement on Auditing Standards No. 47. New York, NY: AICPA.

———. 1983b. *Audit Sampling*. New York, NY: AICPA.

———. 1988. *Consideration of Internal Control in a Financial Statement Audit*. Statement on Auditing Standards No. 55. New York, NY: AICPA.

———. 1997. *Consideration of Fraud in a Financial Statement Audit*. Statement on Auditing Standards No. 82. New York, NY: AICPA.

———. 1999. *Audit Sampling*. New York, NY: AICPA.

———. 2002. *Consideration of Fraud in a Financial Statement Audit*. Statement on Auditing Standards No. 99. New York, NY: AICPA.

Bamber, L., O. Barron, and T. Stober. 1997. Trading volume and different aspects of disagreement coincident with earnings announcements. *The Accounting Review* 72 (4): 575–597.

Bedard, J. 1989. An archival investigation of audit program planning. *Auditing: A Journal of Practice & Theory* 9 (1): 57–71.

———, and K. Johnstone. 2002. Auditors' assessments and responses to earnings management risk and corporate governance risk. Working paper, Northeastern University.

Bell, T., F. Marrs, I. Solomon, and H. Thomas. 1997. *Auditing Organizations Through a Strategic Systems Lens: The KPMG Business Measurement Process*. Montvale, NJ: KPMG Peat Marwick.

———, W. Landsman, and D. Shackelford. 2001. Auditors' perceived business risk and audit fees: Analysis and evidence. *Journal of Accounting Research* 39 (1): 35–44.

Bhattacharya, N. 2001. Investors' trade size and trading responses around earnings announcements: An empirical investigation. *The Accounting Review* 76 (2): 221–244.

Dusenbury, R., J. Reimers, and S. Wheeler. 1996. An empirical study of belief-based and probability-based specifications of audit risk. *Auditing: A Journal of Practice & Theory* 15 (2): 12–28.

Elder, R., and R. Allen. 1998. An empirical investigation of the auditor's decision to project errors. *Auditing: A Journal of Practice & Theory* 17 (2): 71–87.

Froot, K. 1989. Consistent covariance matrix estimation with cross-sectional dependence and heteroskedasticity in financial data. *Journal of Financial and Quantitative Analysis* 14 (September): 333–355.

Hackenbrack, K., and W. Knechel. 1997. Resource allocation decisions in audit engagements. *Contemporary Accounting Research* 14 (Fall): 481–499.

Houghton, C., and J. Fogarty. 1991. Inherent risk. *Auditing: A Journal of Practice & Theory* 10 (1): 1–21.

Houston, R. 1999. The effects of fee pressure and client risk on audit seniors' time budget decisions. *Auditing: A Journal of Practice & Theory* 18 (2): 70–86.

———, M. Peters, and J. Pratt. 1999. The audit risk model, business risk and audit-planning decisions. *The Accounting Review* 74 (3): 281–298.

Johnstone, K., and J. Bedard. 2001. Engagement planning, bid pricing, and client response in the market for initial attest engagements. *The Accounting Review* 76 (2): 199–220.

Kachelmeier, S., and W. Messier, Jr. 1990. An investigation of the influence of a nonstatistical decision aid on auditor sample size decisions. *The Accounting Review* 65 (1): 209–226.

Messier, W., S. Kachelmeier, and K. Jensen. 2001. An experimental assessment of recent professional developments in nonstatistical audit sampling guidance. *Auditing: A Journal of Practice & Theory* 20 (1): 81–96.

Mock, T., and A. Wright. 1993. An exploratory study of auditors' evidential planning judgments. *Auditing: A Journal of Practice & Theory* 12 (2): 39–61.

———, and ———. 1999. Are audit program plans risk-adjusted? *Auditing: A Journal of Practice & Theory* 18 (1): 55–74.

O'Keefe, T., D. Simunic, and M. Stein. 1994. The production of audit services: Evidence from a major public accounting firm. *Journal of Accounting Research* 32 (Autumn): 241–261.

Panel on Audit Effectiveness (PAE). 2000. *The Panel on Audit Effectiveness: Report and Recommendations.* Stamford, CT: Public Oversight Board.

Reimers, J., S. Wheeler, and R. Dusenbury. 1993. The effect of response mode on auditors' control risk assessments. *Auditing: A Journal of Practice & Theory* 12 (2): 62–78.

Shelton, S., O. Whittington, and D. Landsittel. 2001. Auditing firms' fraud risk assessment practices. *Accounting Horizons* 15 (1): 19–34.

Stone, D., and W. Dilla. 1994. When numbers are better than words: The joint effects of response representation and experience on inherent risk judgments. *Auditing: A Journal of Practice & Theory* 13 (2): 1–19.

Waller, W. 1993. Auditors' assessments of inherent and control risk in field settings. *The Accounting Review* 68 (4): 783–803.

Wright, A., and J. Bedard. 2000. Decision processes in audit evidential planning: A multistage investigation. *Auditing: A Journal of Practice & Theory* 19 (1): 123–143.

THE ACCOUNTING REVIEW
Vol. 78, No. 4
2003
pp. 1003–1025

Risk Management in Client Acceptance Decisions

Karla M. Johnstone
University of Wisconsin–Madison

Jean C. Bedard
Northeastern University

ABSTRACT: This paper examines whether risk-management strategies (specifically, the use of specialist personnel and higher billing rates) moderate the effect of risk on client acceptance decisions, thereby assisting auditors in bringing prospective client relationships to acceptable risk/return levels. We propose a conceptual model of the client acceptance decision process, and use archival data on one firm's actual client acceptance decisions to test the model. Our results demonstrate the selective use of risk-management strategies in the client acceptance decision, based on the nature of the risks present for each particular client. Specifically, plans to charge a higher billing rate are associated with a reduction in the negative relationship between client acceptance likelihood and both going-concern risk and public trading status, and plans to assign specialist personnel are associated with a reduction in the negative relationship between client acceptance likelihood and both fraud risk and error risk. Therefore, we provide evidence that while risky clients are less likely to be accepted overall, the application of particular risk-management strategies to particular risks increases the likelihood of accepting such clients.

Keywords: *audit planning; bid pricing; client acceptance; error risk; fraud risk; going-concern risk; risk management.*

Data Availability: *Data used for this paper are derived from a proprietary database.*

We appreciate the comments received during workshops at the University of Wisconsin–Madison, the University of Alberta, the University of South Carolina, Temple University, the 2000 University of Illinois Audit Research Conference, and the 2000 International Symposium on Auditing Research. The comments of Steve Asare, Jere Francis, Yves Gendron, and Rick Tubbs have been especially helpful. We also appreciate the comments of the anonymous reviewers. We thank the firm providing data for this paper, and the advice of firm personnel who assisted us.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted April 2002
Accepted March 2003

I. INTRODUCTION

This paper examines whether risk-management strategies (specifically, the use of specialist personnel and higher billing rates) moderate the effect of risk on client acceptance decisions, thereby assisting auditors in bringing prospective client relationships to acceptable risk/return levels.[1] Client acceptance decisions are increasingly important due to continued fee pressure and litigation risk, which make it essential that audit firms carefully consider the potential benefits and costs of association with prospective clients (*Economist* 1995; MacDonald 1997; Johnstone 2000). Writings of the Public Oversight Board (POB) and practice-oriented research reveal that audit firms are paying increasing attention to client acceptance decisions, by evaluating client-related risks and considering strategies to manage those risks to acceptable levels (POB 2000; Bell et al. 2002). Understanding how risk evaluation and risk management affect client acceptance decisions is crucial because the prospective clients that a firm accepts ultimately determine the nature of the firm's evolving client portfolio (Simunic and Stein 1990; Bedard and Johnstone 2003).

We propose a conceptual model of the client acceptance decision based on prior research, professional standards, and the practices of the participating firm. According to the model, engagement personnel gather information from various sources about the risks associated with a prospective client, and evaluate this information to determine whether the expected balance of risk and return is acceptable. Professional guidance and prior research suggest three main types of risk faced by auditors in this decision context: audit risk, client business risk, and auditor business risk.[2] The model then proposes situations in which risk-management strategies may cost-effectively moderate each client's unique risks.

We test our conceptual model by analyzing actual decisions made by a large professional services firm, using information documented in the firm's client acceptance database.[3] This firm's procedures require that engagement personnel document and evaluate detailed information on prospective clients that they seriously consider (including those that are ultimately rejected), including client risk factors and both specialist personnel assignments and planned engagement fees. We obtained data from these evaluations for all prospective clients considered by the firm during a window of time. We use a logistic regression model of client acceptance decisions, in which the independent variables include measures of the three types of risk faced by auditors in this setting: audit risk (measured as both an error risk and a fraud risk assessment), client business risk (measured as a going-concern risk assessment), and auditor business risk (measured as the client's public versus private trading status). The interactions of these risks with specialist personnel and billing rate risk-management strategies form the basis for hypothesis testing.

Results indicate that although the firm is less likely to accept clients with higher levels of error risk or fraud risk, plans to assign specialist personnel moderate this effect. Similarly, the firm is less likely to accept clients that have going-concern risk or that are publicly traded, but plans to charge a higher billing rate moderate this effect. Therefore, our results demonstrate the selective use of risk-management strategies in the client acceptance decision, based on the nature of the risks present for each particular client. More specifically, the participating firm appears to actively manage fraud and error risk via specialist personnel assignment, and use higher billing rates to manage risks (i.e., going-concern risk and public trading status) that may be more difficult to address via audit investment.

---

[1] "Client acceptance" is the term commonly used by auditors to denote their decision about whether to issue an engagement proposal (e.g., AICPA 1997a).

[2] Audit risk is "the risk that the auditor may unknowingly fail to appropriately modify his opinion on financial statements that are materially misstated" (AICPA 1983, AU 312.02). For purposes of this paper, the term "audit risk" does not include a detection risk component, since detection risk is not independent of the audit. Client business risk is "the risk that the client's economic condition will deteriorate in either the short or long term" (Johnstone 2000, 3; Huss and Jacobs 1991). Auditor business risk is the risk that the audit firm will suffer loss resulting from the engagement (Johnstone 2000, 4; AICPA 1983, AU 312.02; Bell et al. 2002).

[3] A confidentiality agreement precludes identification of the participating firm. These data are subject to the usual limitations associated with proprietary data; we discuss those limitations in the "Limitations and Conclusions" section.

Our paper contributes specifically to research on client acceptance decisions, and to the more general research literatures on risk evaluation, risk adaptation, and audit firm portfolio management. First, it provides a theory and empirically tests a model that describes both the relationship between client acceptance likelihood and various client-acceptance risks, and how the assignment of specialist personnel and higher billing rates are used as risk-management strategies. Extending prior research on client acceptance decisions (e.g., Asare and Knechel 1995; Cohen and Hanno 2000; Johnstone 2000; Asare et al. 2002), this paper provides new insight on the conditions in which audit firms might be willing to accept certain riskier clients. It also provides evidence on the situations in which auditors assign specialist personnel, and the rates of their use, which is important to understanding this contemporary trend in auditing.

Second, this paper extends emerging research on audit firm portfolio management decisions (e.g., Jones and Raghunandan 1998; Francis and Reynolds 2002; Francis and Krishnan 2002; Choi et al. 2003). A primary purpose of this line of research is to investigate whether large audit firm portfolios are becoming more or less risky over time, and the results of this line of research are mixed. An important limitation of these papers is that they are unable to distinguish between purposeful portfolio management decisions of audit firms versus shifts in the financial characteristics of audit firm portfolios that occur due to general shifts in the overall economy; this occurs because studying changes in overall client portfolios over time inherently combines auditors' client acceptance decisions, auditors' client retention decisions, and clients' decisions about whether to retain the services of a particular auditor. In contrast, our paper investigates differences between clients for which the participating firm does or does not wish to provide service, and the conditions under which it might be willing to provide service to higher-risk clients, thereby providing a focused investigation of one aspect of audit firm portfolio management decisions.

Third, this paper reports an analysis of archival data concerning client acceptance decisions, which is important for several reasons. First, most audit firms no longer retain information on rejected clients; without such information, estimating a client acceptance model is impossible. Second, these data reflect decisions made under actual practice conditions, including accountability and economic incentives. Studying decisions made in practice is important in developing realistic theories of auditor behavior (e.g., Gibbins and Jamal 1993; Gibbins and Newton 1994). Third, the data are not subject to limitations commonly associated with behavioral experiments or some archival research (e.g., demand effects, response and sampling bias, or hindsight bias) because they were not collected for research purposes. Rather, the client information was documented contemporaneously with actual client acceptance decisions.

The remainder of this paper is organized as follows. Section II describes prior literature, introduces our conceptual model, and details our hypotheses. Section III describes methods. Sections IV and V contain a discussion of the results and our conclusions.

II. PRIOR LITERATURE AND HYPOTHESES

A Model of the Client Acceptance Decision-Making Process

We model the client acceptance decision as a multistage process (see Figure 1) involving the evaluation of risk and return. The model proposes that auditors evaluate risk and the potential billing rate, determining whether the risk/return is already at an acceptable level, or whether applying additional audit investment or charging a higher billing rate would be necessary to effectively bring the client relationship to an acceptable risk/return level. After making these assessments, the model proposes that auditors either accept the prospective client and begin the proposal bidding process, or end discussions.

Evaluation of Risk/Return in the Client Acceptance Decision

The process of client acceptance begins with practice development activities (e.g., cultivating relationships with prospective clients and disseminating information about the firm), which provide opportunities to evaluate prospective clients (Asare et al. 1994). Once practice development activities provide opportunities for considering potential new clients, the model proposes that the audit firm evaluates the client's risks, evaluates the potential billing rate, and determines whether the risk/return is at an acceptable level. Prior research shows that the evaluation of risk (including audit risk, client business risk, and auditor business risk) is a crucial first step in the client acceptance decision (e.g., Huss and Jacobs 1991; AICPA 1994, 1997a, QC 20.14–20.16). Prior research on auditors' risk responsiveness across a variety of decision-making contexts shows that these basic engagement risks are associated with increased audit costs and greater litigation losses (e.g., Palmrose 1987; O'Keefe et al. 1994; Pratt and Stice 1994; Latham and Linville 1998).[4] In the client acceptance context, the limited amount of behavioral and field research shows that higher levels of audit risk, client business risk, and auditor business risk are associated with reductions in client acceptance likelihood (Asare and Knechel 1995; Cohen and Hanno 2000; Johnstone 2000; Asare et al. 2002).[5] Taken together, these studies provide evidence that riskier clients are less likely to be accepted.

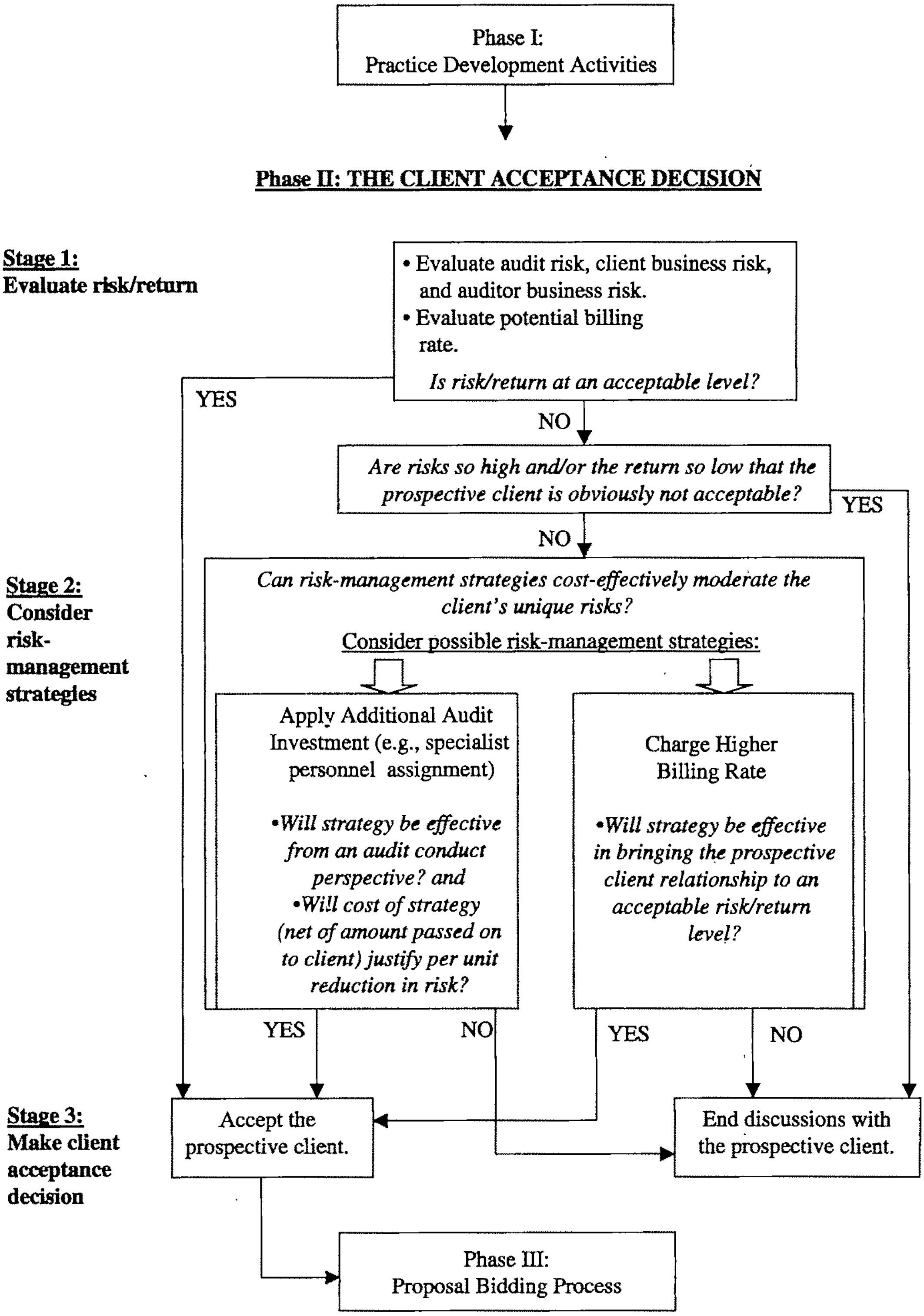
Prior research also demonstrates the importance of potential audit fees as a preliminary consideration in evaluating the risk/return associated with a prospective client. For example, Johnstone (2000) shows that while partners screen clients on the basis of risk, they adjust that "screen" based on their evaluations of the audit firm's risk of financial loss on the

---

[4] Theoretically, audit firms should consider a variety of benefits and costs associated with prospective clients. Benefits include the engagement fee and reputational visibility from serving the client, among others. As modeled by Simunic (1980) and expanded by Houston et al. (2002), the expected total cost of the audit includes the cost of resources used in the engagement (a function of the number of resource units considered necessary to reduce risk to an acceptable level and the anticipated unit cost of those resources), and the present value of expected future losses from the engagement (including litigation due to undetected misstatements and litigation from client business failure or other sources not related to undetected misstatements). Houston et al. (2002) also note that anticipated profits or losses from other future business opportunities may affect a firm's benefit/cost evaluation. Most, if not all, research is necessarily limited in its ability to precisely measure all of these parameters, particularly those involving expected future benefits and costs. In addition, archival research is limited in terms of precise measurement of costs, as such data are generally unavailable from audit firms.

[5] A related stream of archival research using publicly available data on client business risk addresses issues surrounding changes in large audit firm portfolio composition over time (Jones and Raghunandan 1998; Francis and Reynolds 2002; Francis and Krishnan 2002; Choi et al. 2003), and reports mixed results. For instance, Jones and Raghunandan (1998) find that among small manufacturing clients during 1987–1994, audit firm portfolios became less financially risky. Francis and Reynolds (2002) show that during the period 1976–1997 audit firm portfolios became riskier. Francis and Krishnan (2002) show that during the period 1990–1994 audit firm portfolios became riskier, but that during 1995–1997 this increase ceased. Choi et al. (2003) show that audit firm portfolios became riskier during the periods 1985–1989 and 1994–1999, but became less risky during 1990–1994. Therefore, based on client business risk measures, these studies do not provide a clear picture regarding changes in large audit firm portfolio composition over time.

FIGURE 1
A Model of the Client Acceptance Decision-Making Process


Phase I:
Practice Development Activities
Phase II: THE CLIENT ACCEPTANCE DECISION
Stage 1:
Evaluate risk/return
• Evaluate audit risk, client business risk, and auditor business risk.
• Evaluate potential billing rate.
Is risk/return at an acceptable level?
YES
NO
Are risks so high and/or the return so low that the prospective client is obviously not acceptable?
YES
NO
Stage 2:
Consider risk-management strategies
Can risk-management strategies cost-effectively moderate the client's unique risks?
Consider possible risk-management strategies:
Apply Additional Audit Investment (e.g., specialist personnel assignment)
•Will strategy be effective from an audit conduct perspective? and
•Will cost of strategy (net of amount passed on to client) justify per unit reduction in risk?
Charge Higher Billing Rate
•Will strategy be effective in bringing the prospective client relationship to an acceptable risk/return level?
YES
NO
YES
NO
Stage 3:
Make client acceptance decision
Accept the prospective client.
End discussions with the prospective client.
Phase III:
Proposal Bidding Process

engagement—an evaluation dominated by concerns about engagement profitability. Further, Asare et al. (1994) show that potential audit fees can play a key role in an audit firm's decision about whether to accept a prospective client that has approached the firm for possible services.

The client acceptance model in Figure 1 proposes that if the expected risk/return relationship is at an acceptable level, the decision to accept the client is fairly straightforward. The firm begins the bidding process by presenting a bid proposal to the prospective client, perhaps competing with other firms for the engagement (e.g., Johnstone and Bedard 2001; Johnstone et al. 2003). If the expected risk/return relationship is not at an acceptable level, then the model proposes that the audit firm separates the remaining clients into two groups: (1) those that are eliminated from consideration because their risks are so high/and or their potential return is so low that they are obviously not acceptable; and (2) those for which the firm will determine whether available risk-management strategies can cost-effectively moderate the unique set of risks present for each client, thereby bringing the potential client relationships to an acceptable risk/return level. The next section discusses possibilities for the latter of these two groups of prospective clients.

Consideration of Risk-Management Strategies in the Client Acceptance Decision

There are several risk-management strategies that may be applied to an engagement, including personnel-related policies (managing risk through personnel assignment, professional development, and promotion policies), pricing strategies, differential audit activities (e.g., use of tools to facilitate performing, supervising, reviewing, documenting, and communicating audit results), and monitoring policies (managing risk through internal and external peer review) (Robertson and Louwers 1999; Bell et al. 2002).[6] Of these strategies, we study possible risk-moderating effects of specialist personnel assignment and engagement pricing, for two reasons. First, they are the most relevant to client acceptance (i.e., pre-engagement) decision making. Audit firms' bid proposals include engagement pricing and information on personnel to be assigned to engagement teams. Thus, these strategies are likely to be considered and discussed prior to submitting a bid proposal. In contrast, plans for specific audit activities and monitoring strategies are less likely to be shared with clients, and so may not be as thoroughly considered until bids are ultimately accepted by clients. Second, we focus on planned use of specialist personnel due to their importance in reducing risk, according to the current audit approaches of the major auditing firms (e.g., Bell et al. 1997). Further, engagement pricing provides the means to address residual risk not controlled by the assignment of specialists to engagements. Thus, we consider two ways to moderate the client's risks and thereby bring the prospective client relationship to an acceptable risk/return level: a personnel-related strategy that focuses on directly moderating risk by assigning specialist personnel to the engagement team, and a pricing strategy that focuses on indirectly moderating risk by charging higher billing rates to increase the likelihood that the return from the client relationship will be adequate to cover costs.[7] We discuss these risk-management strategies in detail below.

---

[6] In addition to applying these risk-management strategies to an engagement, firms may also choose to eliminate risk by removing certain clients from their portfolios during the client retention decision (e.g., Krishnan and Krishnan 1997; Shu 2000; Bedard and Johnstone 2003).

[7] We acknowledge that billing rates are a proxy for the return that a firm can expect from an engagement. However, as the billing rate increases, it is more likely that the firm will be able to achieve an acceptable risk/return trade-off. The participating firm does not directly measure "anticipated return" (i.e., billing rates minus expected costs) in its client acceptance process.

Specialist personnel assignment as a risk-management strategy. The first risk-management strategy is assignment of specialist personnel, which is encouraged via emerging audit firm methodologies (e.g., Bell et al. 1997; Winograd et al. 2000). While audit firms consider personnel assignment important in risk management, there is very little research that investigates its determinants or its effects on engagement decisions in general (Bedard et al. 1999), and only a few studies are available concerning specialist personnel assignment in the client acceptance context. In a field study, Gendron (2001) describes a situation in which an audit firm made accepting a prospective client contingent on the planned assignment of highly experienced personnel. In a behavioral experiment, Johnstone (2000) expected, but did not find, increased likelihood of accepting a risky client contingent on engagement partners' risk-management strategies, including personnel assignment.[8]

The client acceptance model in Figure 1 proposes that auditors will consider whether assigning specialists will be effective from an audit conduct perspective, and whether the cost of the strategy (net of the amount passed on to the client) will justify the per-unit reduction in risk. The outcome of these considerations is likely to depend on the nature of risks present for any particular prospective client. Prior research has investigated specific differences in specialists' knowledge or decision processes that may be associated with reduction in audit risk. Some of these studies find improved performance in tasks linked to misstatement detection (e.g., Johnson et al. 1991; Wright and Wright 1997), while others find mixed results depending on the metric used or task characteristics (Bedard and Biggs 1991; Solomon et al. 1999). While results vary somewhat, these studies on balance suggest performance gains related to the use of industry specialists, which may serve to moderate audit risk. However, evidence is lacking as to the extent to whether key audit decisions (such as the client acceptance decision, which we study) actually differ due to greater presence of specialists on the engagement team.

In contrast to audit risk, there is little research investigating the likely effects of specialist assignment on reducing client business risk. Theoretically, auditors' understanding of industry practices, economic conditions, and the competitive environment also may help moderate client business risk. The current orientation of several large audit firms toward a business risk audit approach (e.g., Bell et al. 1997) is consistent with the assumption that specialist personnel can improve clients' chances of survival by providing high-quality business advice based on audit findings. In one highly relevant study, Biggs et al. (1993) show that industry specialization of auditors is important in addressing client going-concern problems.

In contrast, assigning specialists to moderate auditor business risk may not be effective from an audit conduct perspective. Audit personnel may specialize in serving public clients (e.g., knowing the necessary reporting requirements for such clients), although we are unaware of research specifically addressing this issue. However, audit firms can be sued on the basis of a decline in a public client's stock price even in the absence of audit failure or client business failure. Thus, we consider that assigning specialist personnel is unlikely to be effective in reducing auditor business risk.

---

[8] The current study overcomes some limitations of Johnstone's (2000) analysis that may have reduced that study's ability to find that personnel assignment moderates risk in the client acceptance context. Particularly, we measure the planned assignment of specialists as a separate strategy, instead of collapsing several strategies into one variable (i.e., personnel assignment, audit pricing, levels of review, and extent of audit testing). In addition, the archival data used in this study reflect decisions made under actual practice conditions that provide incentives for the use of risk management strategies. Our use of a more precise variable and the nature of our data should improve the power of our tests to detect a risk moderation effect of specialist personnel assignment, if it exists.

We propose that assignment of specialist personnel will be associated with a reduction in the negative relationship between risk and client acceptance likelihood for audit risk and client business risk, but not for auditor business risk (see Figure 2, Panel A):

H1a: The planned assignment of specialist personnel will moderate the expected negative relationship between audit risk and client acceptance likelihood (i.e., a positive interaction between audit risk and the assignment of specialist personnel).

H1b: The planned assignment of specialist personnel will moderate the expected negative relationship between client business risk and client acceptance likelihood (i.e., a positive interaction between client business risk and the assignment of specialist personnel).

Charging a higher billing rate as a risk-management strategy. While specialist personnel may reduce the risk of audit failure and litigation exposure, audit firms face residual risk from engagements in which specialists may not be effective from an audit conduct perspective, or are not used because the incremental cost does not justify the risk reduction. For example, as we noted earlier, specialist personnel may be ineffective in reducing auditor business risk from sources such as the client's public trading status. In addition, assigning specialists may not completely moderate client business risk because even high-quality business advice might not solve clients' going-concern problems. Thus, we expect that billing rates will be more likely to moderate the effects of auditor business risk (e.g., publicly traded clients) and client business risk (e.g., going-concern problems) in the client acceptance decision, as these risks are less amenable to investment in the audit than are risks associated with increased likelihood of misstatement.[9]

While auditing standards do not address whether auditors should price risk by charging higher fees, experimental studies find evidence consistent with this risk-management strategy (e.g., Pratt and Stice 1994; Houston et al. 1999). Archival findings vary; some studies detect risk pricing (e.g., Davis et al. 1993; Johnstone and Bedard 2001), while others do not (O'Keefe et al. 1994; Simunic and Stein 1996; Bell et al. 2001). If audit firms consider higher billing rates helpful in managing risk, then the likelihood of accepting riskier clients should increase with the billing rate. While Johnstone's (2000) behavioral experiment finds that the likelihood of accepting risky clients is unaffected by engagement partners' risk-management strategies, this paper extends that work by examining the billing rate as a separate risk-management strategy, and by using data reflecting the incentives existing in actual client acceptance decisions.

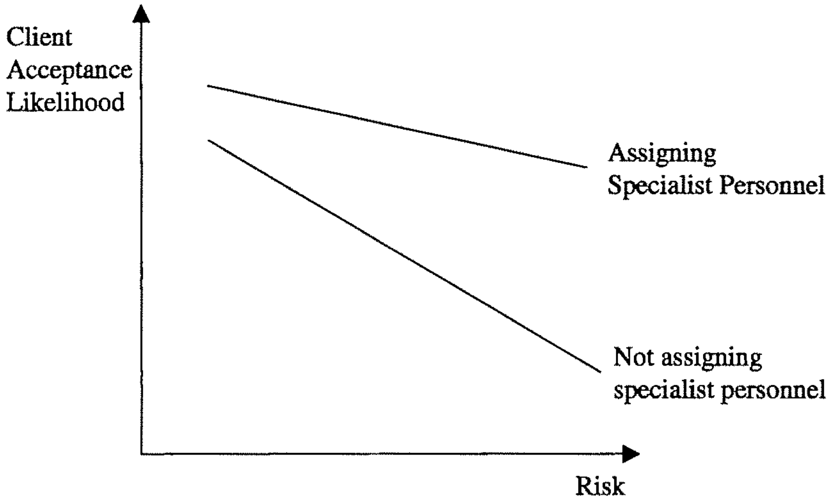
We propose that charging a higher billing rate will be associated with a reduction in the negative relationship between risk and client acceptance likelihood for auditor business risk and client business risk, but not for audit risk (see Figure 2, Panel B):

---

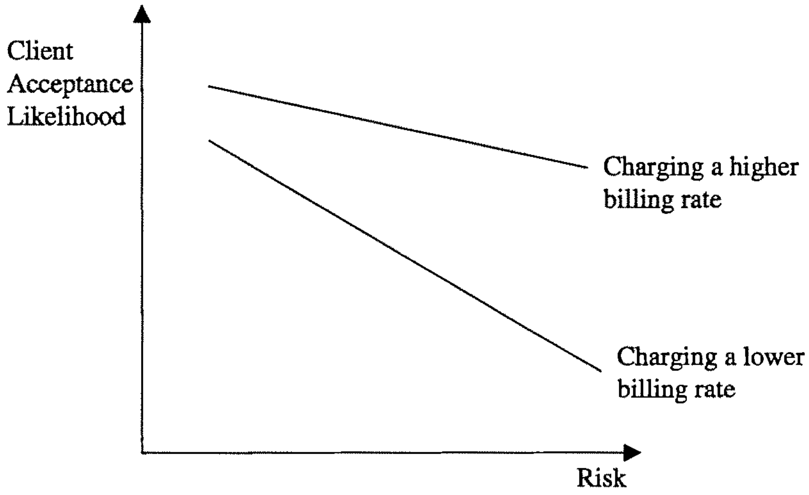
[9] The process of setting audit fees is complex and incompletely understood. Prior research shows that much of the variance in fees is determined by the number of personnel hours, which are in turn explained in large part by client size (e.g., O'Keefe et al. 1994). In examining hourly billing rates, we remove variance associated with numbers of audit hours in order to focus on the revenue available to the firm for each planned hour of engagement activity. Further, engagement fees are not solely determined by cost coverage, as the variance in hourly rates in our database attests. Fees may also be affected by audit firms' demand for certain clients, especially in the initial engagement context that we study. For instance, prior research finds that audit firms "low-ball" initial engagements in order to gain future audit or nonaudit services with presumably higher margins (e.g., Simon and Francis 1988).

FIGURE 2
Illustrations of Proposed Moderating Effects of Risk-management strategies

Panel A: Specialist Personnel Assignment as a Risk Management Strategy (H1)


Client Acceptance Likelihood
Assigning Specialist Personnel
Not assigning specialist personnel
Risk

Panel B: Billing Rate as a Risk Management Strategy (H2)


Client Acceptance Likelihood
Charging a higher billing rate
Charging a lower billing rate
Risk

H2a: A higher billing rate will moderate the expected negative relationship between auditor business risk and client acceptance likelihood (i.e., a positive interaction between auditor business risk and the billing rate).

H2b: A higher billing rate will moderate the expected negative relationship between client business risk and client acceptance likelihood (i.e., a positive interaction between client business risk and the billing rate).

III. METHODS

Sample

Our data comes from evaluations of prospective clients made by the participating firm, comprising information regarding the population of 588 clients that the firm considered between November 1997 and August 1998 (for which its client acceptance decision outcome was known as of the time we obtained the file). We exclude 163 clients due to missing data (primarily financial), leaving a sample of 425 for empirical tests.[10]

The firm's procedures require potential engagement teams to evaluate risk, plan the approach to address risks if the firm undertakes the engagement, and propose a fee. This information, documented in the firm's client acceptance database, forms the basis for its decision on whether to submit a bid proposal for the engagement.[11] The engagement partner, the local office's managing partner, and high-level firm risk-management partners make the client acceptance decision jointly. The final decision is recorded in the firm's database. The firm requires that engagement personnel document information on all prospective clients, including information for prospective clients not ultimately accepted for bid.[12]

Variable Definitions

Dependent and Independent Variables

The dependent variable (*ACCEPTED*) equals 1 if the firm approved issuing an engagement proposal to the prospective client, and 0 otherwise. In terms of independent variables, we measure audit risk using variables indicating the client's risk of fraud and error.[13] Engagement teams complete a checklist containing fraud risk factors (e.g., factors identified in SAS No. 82 [AICPA 1997b], such as management's characteristics and influence over the control environment) and error risk factors (e.g., factors identified in SAS No. 47 [AICPA 1983], such as the client's control environment, the complexity of its transactions, and indications of a client's inability to prepare accurate and timely financial reports). Engagement teams indicate a "yes" or a "no" for the presence of each risk factor. Our fraud risk variable is the sum out of approximately 20 fraud risk factors present for each prospective client (*FRAUD RISK*), and our error risk variable is the sum out of approximately five error risk factors present for each prospective client (*ERROR RISK*).[14]

---

[10] The percentage of clients with missing data is marginally higher for the rejected (37.7 percent) versus the accepted (26.4 percent) client groups (Chi-square = 3.870, p = 0.062). We analyzed the data to determine whether this might be due to a "truncated" completion of the client acceptance form for the rejected clients. We find no case in which the data for a client is discontinued at a certain point during data entry. Rather, some data fields were simply left blank, and those are most commonly found in the financial data section of the form.

[11] Client acceptance decision making at the participating firm involves partners at the local (engagement) level through firm-level risk management partners. The data in the file do not reveal whether particular prospective clients received heightened scrutiny by firm-level risk management partners. Ayers and Kaplan (1998) find that risk management partners make more conservative client acceptance judgments than do engagement partners. The riskiest clients in our sample seem the most likely to receive additional scrutiny by firm-level risk management partners. Ayers and Kaplan's (1998) results imply that the additional scrutiny by these more-conservative partners will bias against finding that risk-management strategies (planned assignment of specialist personnel or higher billing rates) will moderate the effect of risk on client acceptance decisions.

[12] In cases in which a client's risks are so high that the client is obviously not acceptable, the decision to reject the client is likely not documented in the database, so the number of actual client rejections is probably understated. These are the types of clients for which no reasonable amount of risk-management strategies would have likely led to the client being accepted.

[13] All of the risk evaluations are made assuming the absence of the risk-management strategies, i.e., the "pure" risk before applying the strategies.

We measure client business risk using a variable indicating the engagement team's assessment that the prospective client may not remain a going concern (*GC RISK*). Because client business failure is associated with litigation against audit firms (e.g., Carcello and Palmrose 1994), such an assessment should be considered a strong warning sign.[15] We measure auditor business risk using a variable indicating whether the client is publicly traded (*PUBLIC*). The risk associated with public clients is an important component of auditor business risk because lawsuits are most common and the damage awards are greatest for such clients (St. Pierre and Anderson 1984; Lys and Watts 1994). While public clients may be attractive from a practice-development and visibility perspective, prior research on client acceptance decisions shows that audit firms consider public status as a risk factor (Bell et al. 2002).

We measure *SPECIALIST PERSONNEL* using a dichotomous variable indicating whether the engagement team plans to increase its use of specialists on the engagement (i.e., personnel with relatively high industry expertise, functional expertise, and/or high-risk client experience). We measure *BILLING RATE* as the total planned fee divided by the total planned personnel hours. The database does not contain effort estimates by staffing level, so the billing rate is best viewed as an expected average of the rate per hour across all staffing levels and types. Accordingly, if the firm passes along the potentially increased cost of specialist personnel assigned to the client, then their rates would be incorporated into the billing rate. Because billing rates may vary depending on client characteristics such as risk, complexity, and financial condition, we estimate a model of billing rates and save the standardized residuals from that model (*BILLING RATE RESIDUAL*), using that variable as our measure of pricing. Higher values of the *BILLING RATE RESIDUAL* variable represent "excess" billing rates (i.e., positive variation from the billing rate regression line), providing a more precise test of pricing as a risk-management strategy while controlling for client characteristics that may otherwise affect the level of the billing rate.[16]

---

[14] The participating firm has requested that we not fully disclose the specific risk factors. We have included the sum of the factors rather than a weighted combination because we had no *a priori* valid means of identifying which of the factors is more important. To provide further validation for the *FRAUD RISK* and *ERROR RISK* variables, we tested whether an alternative formulation of the factors might be more diagnostic than another set. To do so, we analyzed the frequency of high-risk responses for the factors, and formed new summary measures that include only those fraud or error risk factors that had high-risk answers for ten or more of the clients in the database (i.e., the most common fraud or error risk factors). We re-analyzed the model using these revised measures, and the results are essentially equivalent to those we report.

[15] To provide some validation for the *GC RISK* variable, we tested whether clients with high *GC RISK* were, indeed, financially weaker than clients with low *GC RISK*. We find that clients with high *GC RISK* have higher leverage ($t = -4.774$, $p = 0.000$) and lower return on assets ($t = 2.189$, $p = 0.029$), indicating that they are weaker financially than clients with low *GC RISK*.

[16] We formulate our billing rate model based on the theoretical model of billing rates presented in Bell et al. (2001), which has an adjusted R^2 of 0.040. We proxy for the inherent and business risk elements in their model using the variables *FRAUD RISK*, *ERROR RISK*, and *GC RISK*. We proxy for client size using *ln(REVENUE)*. We proxy for client complexity using the variable *MULTIPLE LOCATIONS*, which equals 1 if the client operates at multiple locations, and 0 otherwise. As in Bell et al. (2001), we control for *PUBLIC* trading status and *LEVERAGE*. We also include industry control variables. Results of the billing rate model are as follows. *BILLING RATE* = 62.84** + 20.59 *FRAUD RISK*** + 3.49 *ERROR RISK* − 2.53 *GC RISK* + 3.04 *ln(REVENUE)*** + 0.62 *MULTIPLE LOCATIONS* + 42.30 *PUBLIC*** − 5.83 *LEVERAGE*** + 0.89 *HIGH TECH* + 3.61 *FINANCIAL SERVICES* + 6.98 *MANUFACTURING* + 9.61 *GOVERNMENT/NONPROFIT* + 6.81 *RETAIL*. The model meets OLS regression assumptions and its adjusted $R^2 = 0.243$. Coefficients that are significant at $p < 0.05$ are represented with **.

Control Variables

We control for client size with the natural log of total revenue (*REVENUE*), and for client financial condition through return on assets (*ROA*) and debt as a percentage of total assets (*LEVERAGE*).[17] The model also includes dichotomous industry control variables to remove any variance associated with differential acceptance rates across industries (*HIGH TECH, GOVERNMENT/NONPROFIT, MANUFACTURING, RETAIL*, and *FINANCIAL SERVICES*).

Model of the Client Acceptance Decision

We test our hypotheses by examining interactive effects of the client acceptance risk variables (*FRAUD RISK, ERROR RISK, GC RISK*, and *PUBLIC*) and the risk-management strategy variables (*SPECIALIST PERSONNEL* and *BILLING RATE RESIDUAL*). We test these hypotheses using a logistic regression whose dependent variable is *ACCEPTED*. Risk moderation implies a positive coefficient on a *Risk Factor* × *Risk-Management Strategy* interaction. The model is as follows:

$$\begin{aligned} ACCEPTED = {} & a_0 + a_1\, FRAUD\ RISK + a_2\, ERROR\ RISK + a_3\, GC\ RISK \\ & + a_4\, PUBLIC + a_5\, SPECIALIST\ PERSONNEL \\ & + a_6\, BILLING\ RATE\ RESIDUAL \\ & + a_{7\text{-}10}\, (RISK\ FACTOR \times SPECIALIST\ PERSONNEL) \\ & + a_{11\text{-}14}\, (RISK\ FACTOR \times BILLING\ RATE\ RESIDUAL) \\ & + a_{15}\, lnREVENUE + a_{16}\, ROA + a_{17}\, LEVERAGE \\ & + a_{18\text{-}22}\, INDUSTRY\ VARIABLES + u. \end{aligned} \tag{1}$$

IV. RESULTS

Descriptive Statistics

Table 1 presents descriptive statistics, including univariate comparisons between accepted- and rejected-client groups. Panel A shows that the firm accepted approximately 90 percent (382/425) of the prospective clients in the model. About 9 percent of clients overall have an indication of high going-concern risk, but this rate is significantly higher in the rejected-client group (30.2 percent) than in the accepted-client group (6.8 percent) (Chi-square = 25.449, p = 0.000). Overall, about 16 percent of clients are publicly traded, and this rate is significantly higher in the rejected-client group (44.2 percent) than in the accepted-client group (12.3 percent) (Chi-square = 29.949, p = 0.000). Engagement teams plan to assign specialist personnel on about 52 percent of prospective clients overall, and this rate is similar across the rejected and the accepted-client groups.[18]

---

[17] We winsorize financial variables at +/−3 standard deviations.

[18] As a side issue, we note that the participating firm assigns specialist personnel to both higher and lower risk clients. For example, the firm assigns specialist personnel to 49 percent of private clients. One reviewer noted that some readers might find this surprising. There is little published research on the issue of audit staffing in general (see Bedard et al. 1999), and no published research (of which we are aware) that directly addresses why firms might assign specialists to lower risk clients. However, audit firms are now stressing the importance of developing specialized knowledge in their personnel (Bell et al. 1997; Winograd et al. 2000), and audit firms compete with each other based upon the assignment of specialized personnel (Johnstone et al. 2003). These observations suggest additional benefits to the assignment of specialist personnel beyond risk management, which seems an important avenue for future research.

TABLE 1
Descriptive Statistics

Panel A: Dichotomous Variables: Number (Percent) = 1

Variable	Variable Description	Full Sample (n = 425)	Accepted Clients (n = 382)	Rejected Clients (n = 43)	Chi-Square
GC RISK	= 1 if indication of going-concern risk in the upcoming engagement; 0 otherwise	39 (9.2%)	26 (6.8 %)	13 (30.2 %)	25.449***
PUBLIC	= 1 for publicly held companies; 0 otherwise	66 (15.5 %)	47 (12.3 %)	19 (44.2 %)	29.949***
SPECIALIST PERSONNEL	= 1 if engagement teams plan increased use of specialists, 0 otherwise;	219 (51.5 %)	198 (51.8 %)	21 (48.8 %)	0.139
HIGH TECH	= 1 for information technology, software, or Internet commerce companies; 0 otherwise	116 (27.3 %)	107 (28.0 %)	9 (20.9 %)	0.976
FINANCIAL SERVICES	= 1 for banks or insurance, investment, or real estate clients; 0 otherwise	50 (11.8 %)	40 (10.5 %)	10 (23.3 %)	6.085**
MANUFACTURING	= 1 for automotive, engineering, construction, or pharmaceutical companies; 0 otherwise	59 (13.9 %)	49 (12.8 %)	10 (23.3 %)	3.516*
GOVERNMENT/ NONPROFIT	= 1 for health care, higher education, nonprofit, or state/local governmental clients; 0 otherwise	86 (20.2 %)	84 (22.0 %)	2 (4.7 %)	7.198***
RETAIL	= 1 for consumer products or retail distribution companies; 0 otherwise	41 (9.6 %)	38 (9.9 %)	3 (6.9 %)	0.391
SERVICE	= 1 for hospitality, entertainment, telecom or transportation companies; 0 otherwise	73 (17.2 %)	63 (16.5 %)	10 (23.3 %)	0.614

(*continued on next page*)

TABLE 1 (continued)

Panel B: Continuous Variables

Variable	Variable Description	Full Sample (n = 425) Mean (Median)	Full Sample (n = 425) Standard Deviation	Full Sample (n = 425) Minimum-Maximum	Accepted Clients (n = 382) Mean (Median)	Rejected Clients (n = 43) Mean (Median)	t-test
FRAUD RISK	Number of fraud risk factors	0.292 (0.000)	0.648	0–5	0.257 (0.000)	0.605 (0.000)	3.383***
ERROR RISK	Number of error risk factors	0.205 (0.000)	0.448	0–3	0.191 (0.000)	0.326 (0.000)	1.871*
BILLING RATE	Planned hourly billing rate	$101 ($100)	$57	$16–$469	$101 ($95)	$108 ($107)	0.834
BILLING RATE RESIDUAL	Standardized residual from billing rate model	0.000 (−0.106)	0.986	−1.910–7.065	0.016 (−0.100)	−0.141 (−0.189)	−0.992
REVENUE	Total revenue	$314,258 ($8,077)	$1,311,771	$1–$8,333,000	$297,187 ($8,503)	$465,918 ($4,743)	0.799
ROA	Net income ÷ Total assets	−0.208 (0.020)	2.219	−20.57–29.61	−0.179 (0.027)	−0.464 (−0.599)	−0.799
LEVERAGE	Total liabilities ÷ Total assets	0.744 (0.575)	0.982	0–6.82	0.695 (0.559)	1.184 (0.777)	3.131***

---

*, **, *** Indicate significant differences in characteristics between prospective clients that the firm accepted or rejected, at the 10 percent, 5 percent and 1 percent levels, respectively, in two-tailed tests.

Descriptive statistics for dichotomous control variables show that the highest industry concentrations in the sample are in high-technology (27.3 percent), government/nonprofit (20.2 percent), and service (17.2 percent) industries. The univariate tests show that the participating firm is more likely to accept clients in the government/nonprofit industry (Chi-square = 7.198, p = 0.007), and is less likely to accept clients in the financial services (Chi-square = 6.085, p = 0.014) and manufacturing (Chi-square = 3.516, p = 0.061) industries.

Panel B of Table 1 shows that the mean number of fraud risk factors is 0.292, and that the rejected-client group has a higher average number of fraud risk factors (0.605) than the accepted-client group (0.257) (t = 3.383, p = 0.001).[19] The mean number of error risk factors is 0.205, and the rejected-client group has a marginally higher average number of error risk factors (0.326) than the accepted-client group (0.191) (t = 1.871, p = 0.062). The mean (median) billing rate is \$101 (\$100), and the mean (median) client size in terms of revenue is about \$314 (\$8) million.[20] The mean (median) *ROA* is −0.208 (0.020), which is similar across the rejected- and the accepted-client groups. The mean (median) *LEVERAGE* is 0.744 (0.575), and the rejected-client group has higher *LEVERAGE* (1.184) than the accepted-client group (0.695) (t = 3.131, p = 0.002).[21]

Table 2 presents correlations among the variables, showing that the firm is less likely to accept clients with fraud risk, error risk, going-concern risk, and those clients that are publicly traded or highly leveraged.

Model of the Client Acceptance Decision

Table 3 presents results of Model (1), a logistic regression model of client acceptance likelihood on risk factors, risk-management strategies, interactions between the risk factors and risk-management strategies, and control variables. The model is significant, with a Wald Chi-square statistic of 81.085, and pseudo-R^2 of 36.1 percent.

Model results show that client acceptance likelihood is negatively associated with *FRAUD RISK* (Chi-square = 5.050, p = 0.013), *ERROR RISK* (Chi-square = 1.981, p = 0.079), *GC RISK* (Chi-square = 8.339, p = 0.002) and *PUBLIC* (Chi-square = 13.401, p = 0.000). These findings are consistent with prior behavioral experiments showing that auditors are less likely to accept riskier clients (e.g., Johnstone 2000), and with audit firms' current client acceptance data collection strategies (whereby, for example, public trading status is viewed as a risk factor [Bell et al. 2002]).

---

[19] The relatively low mean number of fraud risk factors for clients in our sample is consistent with emerging research on the prevalence of fraud risk factors. For example, Graham and Bedard (2003) report an average of 1.85 fraud risk factors for clients in their sample, and Mock and Turner (2001) report an average of 2.9 fraud risk factors for clients in their sample.

[20] Investigating those clients with very low billing rates, we find that they are all private clients that the firm decided to accept. We find that the lowest decile of billing rates for these clients ranges from \$16 to \$40, while the lowest billing rate for rejected clients is \$46. All of the 39 prospective clients with billing rates below \$40 that were accepted by the firm are private. Of those, only one has a fraud risk factor, three are considered going-concern risks, and 34 are in the high-technology industry. These findings suggest differential acceptance strategies in the public- and private-client markets, particularly for high-technology clients assessed as having relatively low risk. Thus, during the 1997–1998 period of this study, the firm may have been "investing" in such clients by offering very low first-year billing rates. Investigating those clients with very low revenue, we find that they are primarily private, high-technology clients that are highly leveraged (i.e., high-technology start-ups).

[21] To provide a comparison between our sample and the market as a whole during this period, we calculated descriptive statistics from the Compustat database on our financial control variables. Our sample is fairly similar to other public companies on these variables. For example, in 1997, mean (median) leverage for Compustat companies is 0.896 (0.569), and in 1998 it is 0.896 (0.590). In 1997, mean (median) return on assets is −0.167 (−0.017), and in 1998 it is −0.298 (0.011).

TABLE 2
Correlation Matrix

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
1. *ACCEPTED*	1.00															
2. *FRAUD RISK*	−0.16***	1.00														
3. *ERROR RISK*	−0.09*	0.25***	1.00													
4. *GC RISK*	−0.25***	0.10**	0.20***	1.00												
5. *PUBLIC*	−0.27***	0.28***	0.11**	−0.01	1.00											
6. *BILLING RATE*	−0.04	0.31***	0.14***	−0.02	0.38***	1.00										
7. *BILLING RATE RESIDUAL*	0.05	0.00	0.00	0.00	0.00	0.86***	1.00									
8. *SPECIALIST PERSONNEL*	0.02	0.05	0.07	−0.01	0.10**	−0.10*	−0.13**	1.00								
9. *lnREVENUE*	0.04	0.06	0.15***	−0.01	0.21***	0.32**	0.00	−0.07	1.00							
10. *ROA*	0.04	0.03	−0.01	−0.11**	−0.01	0.06	−0.02	−0.09*	0.12**	1.00						
11. *LEVERAGE*	−0.15***	0.13***	0.05	0.23***	−0.04	−0.13***	0.00	0.07	−0.20***	−0.37**	1.00					
12. *HIGH TECH*	0.05	−0.03	−0.04	0.08	0.00	−0.13***	0.00	0.22***	0.33***	−0.19***	0.17***	1.00				
13. *FIN. SVCS.*	−0.12**	0.04	0.08	−0.02	0.13***	0.06	0.00	0.05	0.05	0.04	−0.03	−0.22***	1.00			
14. *MFG.*	−0.09*	0.06	0.12**	0.01	0.02	0.06	0.00	−0.14***	0.11**	0.03	−0.01	−0.25***	−0.15***	1.00		
15. *GOVT/ NONPROFIT*	0.13***	−0.12**	−0.14***	−0.10**	−0.15***	0.02	0.00	−0.01	0.13***	0.09*	−0.13***	−0.31***	−0.18***	−0.20***	1.00	
16. *RETAIL*	0.03	0.03	−0.04	−0.05	−0.03	0.04	0.00	−0.15***	0.11**	0.05	−0.03	−0.20***	−0.12**	−0.13***	−0.17***	1.00

*, **, *** Indicate that the estimated coefficients are significantly different from zero at the 10 percent, 5 percent and 1 percent levels, respectively, in two-tailed tests.
This table reports Pearson correlations among independent variables.

TABLE 3
Logistic Regression of Client Acceptance Decisions

$$
\begin{aligned}
\mathit{ACCEPTED} = {} & a_0 + a_1\,\mathit{FRAUD\ RISK} + a_2\,\mathit{ERROR\ RISK} + a_3\,\mathit{GC\ RISK} \\
& + a_4\,\mathit{PUBLIC} + a_5\,\mathit{SPECIALIST\ PERSONNEL} \\
& + a_6\,\mathit{BILLING\ RATE\ RESIDUAL} \\
& + a_{7\text{-}10}\,(\mathit{RISK\ FACTOR} \times \mathit{SPECIALIST\ PERSONNEL}) \\
& + a_{11\text{-}14}\,(\mathit{RISK\ FACTOR} \times \mathit{BILLING\ RATE\ RESIDUAL}) \\
& + a_{15}\,\mathit{ln\ REVENUE} + a_{16}\,\mathit{ROA} + a_{17}\,\mathit{LEVERAGE} \\
& + a_{18\text{-}22}\,\mathit{INDUSTRY\ VARIABLES} + u \qquad (1)
\end{aligned}
$$

	Predicted Sign	B (Chi-square)
Intercept		1.001 (0.794)
FRAUD RISK	−	−0.979 (5.050)**
ERROR RISK	−	−0.804 (1.981)*
GC RISK	−	−2.140 (8.339)***
PUBLIC	−	−2.447 (13.401)***
SPECIALIST PERSONNEL		−0.964 (2.312)
BILLING RATE RESIDUAL		−0.507 (3.676)*
FRAUD RISK × *SPECIALIST PERSONNEL* (H1a)	+	1.085 (3.012)**
ERROR RISK × *SPECIALIST PERSONNEL* (H1a)	+	1.981 (4.394)**
GC RISK × *SPECIALIST PERSONNEL* (H1b)	+	0.280 (0.072)
PUBLIC × *SPECIALIST PERSONNEL*	+	0.887 (1.017)
FRAUD RISK × *BILLING RATE RESIDUAL*	+	0.098 (0.171)
ERROR RISK × *BILLING RATE RESIDUAL*	+	0.285 (0.235)
GC RISK × *BILLING RATE RESIDUAL* (H2b)	+	1.034 (2.145)*
PUBLIC × *BILLING RATE RESIDUAL* (H2a)	+	0.798 (3.756)**
Control Variables		
ln(REVENUE)		0.093 (3.100)*
ROA		−0.038 (0.177)
LEVERAGE		−0.256 (2.345)
HIGH TECH		1.024 (2.661)
FINANCIAL SERVICES		−0.562 (0.907)
MANUFACTURING		−0.585 (0.977)
GOVERNMENT/NONPROFIT		1.604 (2.819)*
RETAIL		0.152 (0.038)

n = 425; Model Chi-Square = 81.085;
Nagelkerke $R^2 = 0.361$

---

*, **, *** Indicate significant coefficients at the 10 percent, 5 percent, and 1 percent levels, respectively. We report one-tailed probabilities for tests of directional expectations, and report two-tailed probabilities for other tests. This table presents results of a logistic regression model examining the effect of client acceptance risks and risk-management strategies on the firm's client acceptance decisions.

Our research hypotheses relate to the interactive effects of risk factors and risk-management strategies. Specifically, we expect that the assignment of specialist personnel (H1) or a higher billing rate (H2) will be associated with a reduction in the negative relationship between risk and client acceptance likelihood. Regarding H1, results of Model (1) show that the *FRAUD RISK* × *SPECIALIST PERSONNEL* interaction is positive and significant (Chi-square = 3.012, p = 0.042), as is the *ERROR RISK* × *SPECIALIST PERSONNEL* interaction (Chi-square = 4.394, p = 0.018), but the *GC RISK* × *SPECIALIST PERSONNEL* interaction is not significant. Thus, our findings are consistent with H1a, but not with H1b.

Interpreting the magnitudes of effects related to *FRAUD RISK* and its interaction with *SPECIALIST PERSONNEL*, we find that the probability of accepting a client with one fraud risk factor is 50.5 percent, whereas the probability of accepting a client with two fraud risk factors is 27.7 percent. The coefficient of the *FRAUD RISK* × *SPECIALIST PERSONNEL* interaction implies that the probability of accepting a client with two fraud risk factors rises to 56.2 percent when specialists are assigned.[22] Similarly, the probability of accepting a client with one error risk factor is 54.9 percent, whereas the probability of accepting a client with two error risk factors is 35.3 percent. The coefficient of the *ERROR RISK* × *SPECIALIST PERSONNEL* interaction implies that the probability of accepting a client with two error risk factors rises to 91.6 percent when specialists are assigned.

Regarding H2, results of Model (1) show that the *PUBLIC* × *BILLING RATE RESIDUAL* interaction is positive and significant (Chi-square = 3.756, p = 0.027) and that the *GC RISK* × *BILLING RATE RESIDUAL* interaction is positive and marginally significant (Chi-square = 2.145, p = 0.072). However, the *FRAUD RISK* × *BILLING RATE RESIDUAL* and the *ERROR RISK* × *BILLING RATE RESIDUAL* interactions are not significant.[23] Thus, our findings are consistent with both H2a and H2b.

Interpreting these effects, the coefficient on the main effect for *PUBLIC* implies that the probability of accepting a client that is not publicly traded is 73.1 percent, while the probability of accepting a client that is publicly traded is 19.1 percent. The coefficient of the *PUBLIC* × *BILLING RATE RESIDUAL* interaction implies that the probability of accepting a publicly traded client rises to 29.7 percent when the billing rate is two standard deviations higher than average. For *GC RISK*, the probability of accepting a client without going-concern risk is 73.1 percent, while the probability of accepting a client identified as a going-concern risk is 24.2 percent. The coefficient of the *GC RISK* × *BILLING RATE RESIDUAL* interaction implies that the probability of accepting a client with going-concern risk rises to 47.9 percent when the billing rate is two standard deviations higher than average.

The control variables in Model (1) are generally not significant, except that the firm is somewhat more likely to accept larger clients (Chi-square = 3.100, p = 0.078), and those in the *GOVERNMENT/NONPROFIT* industry (Chi-square = 2.819, p = 0.093). The financial control variables (i.e., *LEVERAGE* and *ROA*) may be insignificant in the model because the risk inherent in those variables might already be captured via the assessments of *FRAUD RISK*, *ERROR RISK*, or *GC RISK*. For example, prior research shows that auditors' judgments of audit risk are related to client financial health (e.g., Johnstone 2000; Graham and Bedard 2003). In order to investigate this possibility, we estimated regression models with each risk measure (i.e., *FRAUD RISK*, *ERROR RISK*, and *GC RISK*) as dependent variables, and the financial control variables as independent variables. The results show that *LEVERAGE* is significantly associated with *FRAUD RISK* and *GC RISK*, but not *ERROR RISK*, and that *ROA* is not associated with any of the risk assessments. Therefore, *LEVERAGE* may lose significance because some of its effect is already captured via variance in the firm's risk assessments, whereas *ROA* appears unrelated to the likelihood of accepting a prospective client.

---

[22] The reported probabilities were computed as follows (for further information on interpreting coefficients of interaction terms in logistic regression, see Jaccard [2001]). For a client with one fraud risk factor, the log odds = 1.001 − 0.979 (1) = 0.022; the odds = $e^{0.022}$ = 1.022; and the probability of acceptance = 1.022/(1+1.022) = 0.505. For a client with two fraud risk factors, the log odds = 1.001 − 0.979 (2) = −0.957; the odds = $e^{-0.957}$ = 0.384; and the probability of acceptance = 0.384/(1 + 0.384) = 0.277. For a client with two fraud risk factors and with specialists assigned, the log odds = 1.001 − 0.979 (2) −0.964 (1) + 1.085 (2) = 0.249; the odds = $e^{0.249}$ = 1.283; and the probability of acceptance = 1.283/(1 + 1.283) = 0.562. For a client with two fraud risk factors and with no specialists assigned, the log odds = 1.001 − 0.979 (2) −0.964 (0) + 1.085 (0) = −0.957; the odds = $e^{-0.957}$ = 0.384; and the probability of acceptance = 0.384/(1 + 0.384) = 0.277. Other interpretations of coefficients in the text are calculated similarly.

[23] We note that the coefficient on the *BILLING RATE RESIDUAL* is marginally negative, which suggests that the firm is more likely to accept lower-risk clients with fees lower than "standard." The significant negative effect of billing rate is sensitive to the inclusion of the 39 private, high-technology clients with low billing rates that were all accepted by the firm (as discussed in footnote 20). When we remove those clients from the analysis, the *BILLING RATE RESIDUAL* is no longer significant (B = −0.292, p = 0.387), and inferences regarding hypothesis tests are unchanged.

Overall, the results are consistent with our conceptual model. The firm evaluates audit risk, client business risk, and auditor business risk, and reduces client acceptance likelihood in the presence of those risks. It also selectively uses personnel assignment and pricing strategies to moderate the effects of those risks on client acceptance likelihood, thereby making it possible for the firm to bring the risk/return associated with some prospective clients to an acceptable level.

V. LIMITATIONS AND CONCLUSIONS

Limitations

Prior to discussing our results, we address several limitations of the study, each of which provides opportunities for further research. First, the data represent decisions made by a single firm, and our analysis is constrained to the variables that the firm believed important enough to include in its database. Therefore, understanding the generalizability of our results awaits disclosure of similar data by other firms. Second, engagement teams could have an incentive to make the client appear more attractive to improve the chances of acceptance. However, we assume that the effects of any incentives for misrepresentation are offset by the numerous legal, professional, and reputational costs associated with misrepresenting the known facts about a client. If engagement teams did attempt to make the client appear more attractive, then those clients would be represented as "low risk" in the database, so it seems unlikely that the engagement teams would suggest accompanying risk-management strategies. The variation in risk present in the database would be tempered by such behavior, making it more difficult to detect significant effects.

Third, our model proposes that auditors will apply additional audit investment to moderate client acceptance risks when they expect that such a strategy will be effective from an audit conduct perspective, and when its cost justifies the anticipated reduction in risk. Our theoretical focus is on the audit effectiveness aspect of this risk-management strategy, and our empirical data do not allow us to determine whether the cost of assigning specialists is justified on the basis of a per unit reduction in risk. Therefore, we encourage future research that investigates this aspect of risk management. Fourth, our model accounts for about 36 percent of the variance in client acceptance decisions, so future research could investigate other potential explanatory factors in this decision (e.g., audit firm strategy, audit firm market position, market-level competition, or client-related risks that we do not include in our model).

Conclusions

This paper examines evaluations of prospective clients derived from a participating firm's client acceptance decisions to test whether risk-management strategies moderate the effect of risk on client acceptance decisions. We find that the intent to use specialist personnel moderates the negative relationship between audit risk (i.e., fraud and error risk) and client acceptance likelihood. The risk-moderating effect of specialized personnel on fraud risk and error risk is consistent with prior behavioral research (e.g., Wright and Wright 1997; Solomon et al. 1999), showing that industry specialists outperform nonspecialists on some audit tasks within their industry domain. We infer from this result that the firm considers that assigning such personnel alters the risk-return balance to an extent sufficient to warrant acceptance of some clients that would otherwise be unacceptable. This finding is important in that it demonstrates that outcomes of a key engagement decision vary on the basis of specialist personnel assignment. However, it leads to a crucial question: Is the implied strategy effective? We cannot address this question directly, as our data represent the *intention* to apply industry specialists to new clients. Further research should investigate whether actual engagement outcomes differ. Furthermore, if the use of industry specialists does in fact reduce audit risk, then what specific actions by specialists are most effective in reducing audit risk? Researchers have investigated differences between specialists and nonspecialists using a variety of auditing tasks and decision metrics, and findings vary. Our understanding of effects of specialization on audit quality would benefit from further research that systematically investigates specialists' application of their knowledge across audit tasks, focusing on where gains from use of specialists are most likely to occur, how those gains can be leveraged by firms, and factors inhibiting them.

Contrary to our expectation, we find no evidence that specialist personnel moderate the negative relationship between client business risk (going-concern risk) and client acceptance likelihood. This result is inconsistent with business risk audit approaches, implying that the firm does not consider that specialists are effective in applying their knowledge to help clients out of financial difficulty. One explanation for our failure to find a risk-moderating effect of specialist personnel on client business risk is that the clients in our sample that are identified as having going-concern risk are very troubled. The firm apparently chooses not to be associated with these prospective clients regardless of engagement team composition. It is quite possible that specialist personnel assignment may affect outcomes of other audit decisions for continuing clients with which the firm already has a relationship. For existing clients, more is known about the client's economic conditions and business process. As information asymmetry is reduced, the firm is in a better position to assess whether specialists can indeed make a difference in assisting client business turnaround. Further research should investigate this issue.

We also find that charging higher billing rates moderates the negative relationship between both going-concern risk and public trading status and client acceptance likelihood, but does not moderate the negative relationship between audit risk and client acceptance likelihood. Together, these results suggest that the firm considers specialist personnel to be effective in detecting and addressing fraud and error risks, and considers higher billing rates to be effective in addressing the risks associated with client financial failure or public trading status. This difference in risk-management focus may be related to the fact that audit firms are often sued on the basis of client bankruptcy or stock price declines even in the absence of an audit failure, and helps to explain the mixed results of prior studies that demonstrate the differential application of fee premia to some risks but not others (Houston et al. 1999, 2002).[24] Future research would be helpful in specifically investigating how higher billing rates are used to moderate risks on high-risk engagements, and/or whether and how such monies could be pooled to provide funds to moderate risk at the audit firm portfolio level.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 1983. *Audit Risk and Materiality in Conducting an Audit.* Statement on Auditing Standards No. 47. New York, NY: AICPA.

———. 1994. *Acceptance and Continuation of Audit Clients.* Practice Alert No. 94-3. New York, NY: AICPA.

———. 1997a. System of quality control for a CPA firm's accounting and auditing practice:. QC 20.14–20.16. New York, NY: AICPA.

———. 1997b. *Consideration of Fraud in a Financial Statement Audit.* Statement on Auditing Standards No. 82. New York, NY: AICPA.

Asare, S., K. Hackenbrack, and W. R. Knechel. 1994. Client acceptance and continuation decisions. In *Proceedings of the 1994 Deloitte & Touche/University of Kansas Symposium on Auditing Problems,* edited by R. P. Srivastava, 163–178. Lawrence, KS: University of Kansas.

———, and W. R. Knechel. 1995. Termination of information evaluation in auditing. *Journal of Behavioral Decision Making* 8: 21–31.

———, J. R. Cohen, and G. Trompeter. 2002. The effect of management integrity and non-audit services on client acceptance and staffing decisions. Working paper, University of Florida.

Ayers, S., and S. Kaplan. 1998. Potential differences between engagement and risk review partners and their effect on client acceptance judgments. *Accounting Horizons* 12 (June): 139–153.

Bedard, J. C., and S. F. Biggs. 1991. The effect of domain-specific experience on evaluation of management representations in analytical procedures. *Auditing: A Journal of Practice & Theory* 10 (Supplement): 77–90.

———, T. Mock, and A. Wright. 1999. Evidential planning in auditing: A review of the empirical research. *Journal of Accounting Literature* (18): 96–142.

———, and K. M. Johnstone. 2003. Audit firm portfolio management decisions. Working paper, Northeastern University.

Bell, T. B., F. Marrs, I. Solomon, and H. Thomas. 1997. *Auditing Organizations through a Strategic-Systems Lens.* Montvale, NJ: KPMG Peat Marwick LLP.

———, W. R. Landsman, and D. A. Shackelford. 2001. Auditors' perceived business risk and audit fees: Analysis and evidence. *Journal of Accounting Research* 39 (June): 35–43.

———, J. C. Bedard, K. M. Johnstone, and E. F. Smith. 2002. KRisk^{SM}: A computerized decision aid for client acceptance and continuance risk assessments. *Auditing: A Journal of Practice & Theory* (September): 97–113.

---

[24] The risk-moderating effect of specialized personnel on fraud risk differs from the findings in Houston et al. (1999), who find that the effect of litigation risk on planned audit evidence collection and planned audit fees is greater in their irregularity condition than in their error condition. The authors state that this finding suggests that "auditors perceive that they can (cannot) adequately control the risks introduced by an error (irregularity) by increasing the audit investment" (Houston et al. 1999, 294). Two important differences exist between our study and Houston et al. (1999), which limit comparison between the two studies. First, the dependent variables in the two studies differ, with our study using the likelihood of accepting a potential client as the dependent variable, and their study using audit effort and fees as the dependent variables. As such, the task and focus of the two studies differs. Second, we find that auditors are less likely to accept clients as fraud risk increases. However, Houston et al. (1999, 295) note that "auditors participating in pilot testing of the case indicated that they would continue to pursue the client in the irregularity condition, which was intended to represent an irregularity, but not one so severe that the auditors would not pursue the client." This may be due to the fact that Houston et al. (1999) operationalize the irregularity condition as an intentional inconsistency with GAAP that the auditors have already discovered. As such, the experimental design in Houston et al. (1999) does not allow for the type of test of heightened fraud risk that our archival data allows.

Biggs, S. F., M. Selfridge, and G. R. Krupka. 1993. A computational model of auditor knowledge and reasoning processes in the going-concern judgment. *Auditing: A Journal of Practice & Theory* 12 (Supplement): 82–99.

Carcello, J., and Z-V. Palmrose. 1994. Auditor litigation and modified reporting on bankrupt clients. *Journal of Accounting Research* 32: 1–30.

Choi, J. H., R. Doogar, and A. Ganguly. 2003. Audit firm client portfolios and changes in audit liability regimes: Evidence from the U.S. audit market. Working paper, University of Illinois.

Cohen, J. R., and D. M. Hanno. 2000. Auditors' consideration of corporate governance and management control philosophy in preplanning and planning judgments. *Auditing: A Journal of Practice & Theory* 19 (Fall): 133–146.

Davis, L., D. Ricchiute, and G. Trompeter. 1993. Audit effort, audit fees, and the provision of nonaudit services to audit clients. *The Accounting Review* 68 (January): 135–150.

Economist. 1995. Accountancy firms: A glimmer of hope. London (April 1): 62.

Francis, J. R., and J. K. Reynolds. 2002. Do large accounting firms screen out risky clients? Working paper, University of Missouri–Columbia.

———, and J. Krishnan. 2002. Evidence on accounting firm risk-management strategies before and after the Private Securities Litigation Reform Act of 1995. *Asia-Pacific Journal of Accounting and Economics* 9 (2): 135–157.

Gendron, Y. 2001. The difficult client-acceptance decision in Canadian audit firms: A field investigation. *Contemporary Accounting Research* 18 (2): 283–310.

Gibbins, M., and K. Jamal. 1993. Problem-centered research and knowledge-based theory in the professional accounting setting. *Accounting, Organizations and Society* 18 (5): 451–466.

———, and J. D. Newton. 1994. An empirical exploration of complex accountability in public accounting. *Journal of Accounting Research* 32 (2): 165–186.

Graham, L. E., and J. C. Bedard. 2003. Fraud risk and audit planning. *International Journal of Auditing* (March): 55–70.

Houston, R., M. Peters, and J. Pratt. 1999. The audit risk model, business risk, and audit planning decisions. *The Accounting Review* 74 (July): 281–298.

———, ———, and ———. 2002. How auditors price business risk: A framework and experiment. Working paper, University of Maryland.

Huss, H. F., and F. A. Jacobs. 1991. Risk containment: Exploring auditor decisions in the engagement process. *Auditing: A Journal of Practice & Theory* (Fall): 16–32.

Jaccard, J. 2001. *Interaction Effects in Logistic Regression.* Sage University Papers Series: Quantitative Applications in the Social Sciences (Series/Number 07-135). Thousand Oaks, CA: Sage Publications.

Johnson, P. E., K. Jamal, and R. G. Berryman. 1991. Effects of framing on auditor decisions. *Organizational Behavior and Human Decision Processes* 50: 75–105.

Johnstone, K. M. 2000. Client acceptance decisions: Simultaneous effects of client business risk, audit risk, auditor business risk, and risk adaptation. *Auditing: A Journal of Practice & Theory* (Spring): 1–25.

———, and J. C. Bedard. 2001. Engagement planning, bid pricing, and client response in the market for initial attest engagements. *The Accounting Review* 76 (April): 199–220.

———, ———, and M. L. Ettredge. 2003. The effect of competitive bidding on engagement planning and pricing. *Contemporary Accounting Research* (forthcoming).

Jones, F. L., and K. Raghunandan. 1998. Client risk and recent changes in the market for audit services. *Journal of Accounting and Public Policy* (17): 169–181.

Krishnan, J., and J. Krishnan. 1997. Litigation risk and auditor resignations. *The Accounting Review* 72 (October): 539–560.

Latham, C., and M. Linville. 1998. A review of the literature in audit litigation. *Journal of Accounting Literature* 17: 175–213.

Lys, T., and R. Watts. 1994. Lawsuits against auditors. *Journal of Accounting Research* 32 (Supplement): 65–93.

MacDonald, E. 1997. More accounting firms are dumping risky clients. *Wall Street Journal* (April 25): Section 3, p. 2.

Mock, T., and J. L. Turner. 2001. An archival study of audit fraud risk assessments made under SAS No. 82. Working paper, University of Southern California.

O'Keefe, T., D. Simunic, and M. Stein. 1994. The production of audit services: Evidence from a major public accounting firm. *Journal of Accounting Research* 32 (2): 241–261.

Palmrose, Z-V. 1987. Litigation and independent auditors: The role of business failures and management fraud. *Auditing: A Journal of Practice & Theory* 7: 90–103.

Pratt, J., and J. Stice. 1994. The effects of client characteristics on auditor litigation risk judgments, required audit evidence, and recommended audit fees. *The Accounting Review* 69 (October): 639–655.

Public Oversight Board (POB), The Panel on Audit Effectiveness. 2000. *Report and Recommendations*. Stamford, CT: POB.

Robertson, J. C., and T. J. Louwers. 1999. *Auditing*. Ninth edition. Boston, MA: Irwin/McGraw-Hill.

St. Pierre, K., and J. Anderson. 1984. An analysis of the factors associated with lawsuits against public accountants. *The Accounting Review* 59 (April): 242–263.

Shu, S. Z. 2000. Auditor resignations: Clientele effects and legal liability. *Journal of Accounting and Economics* 29 (2): 173–205.

Simon, D. T., and J. R. Francis. 1988. The effects of auditor change on audit fees: Tests of price cutting and price recovery. *The Accounting Review* (April): 255–269.

Simunic, D. 1980. The pricing of audit services: Theory and evidence. *Journal of Accounting Research* 18 (1): 161–190.

———, and M. Stein. 1990. Audit risk in a client portfolio context. *Contemporary Accounting Research* 6 (Spring): 329–343.

———, and ———. 1996. The impact of litigation risk on audit pricing: A review of the economics and the evidence. *Auditing: A Journal of Practice & Theory* 15 (Supplement): 119–134.

Solomon, I., M. Shields, and O. Whittington. 1999. What do industry-specialist auditors know? *Journal of Accounting Research* 37 (1): 191–208.

Winograd, B. N., J. S. Gerson, and B. L. Berlin. 2000. Audit practices of PricewaterhouseCoopers. *Auditing: A Journal of Practice & Theory* (Fall): 175–182.

Wright, S., and A. Wright. 1997. The effect of industry experience on hypothesis generation and audit planning decisions. *Behavioral Research in Accounting* 9: 273–294.

THE ACCOUNTING REVIEW
Vol. 78, No. 4
2003
pp. 1027–1047

The Value Relevance of the Foreign Translation Adjustment

Henock Louis
Pennsylvania State University

ABSTRACT: The study presents an economic analysis of the foreign translation adjustment and empirically examines the association between change in firm value and the foreign translation adjustment for a sample of manufacturing firms. The study shows that, for firms in the manufacturing sector, the translation adjustment is associated with a loss of value instead of an increase in value. This result stems from the fact that, for firms in the manufacturing sector, the accounting rules governing foreign currency translations generally produce results opposite to the economic effects of exchange rate changes.

I. INTRODUCTION

Proponents of comprehensive income maintain that comprehensive income "identifies all (recognized) sources of value created in one number as a measure of value added" (cf. AAA FASC 1997). In this study, I present an economic analysis of one of the major components of comprehensive income, the foreign translation adjustment. The analysis suggests that, in general, the translation adjustment is not a source of value added for firms in the manufacturing sector because the accounting rules governing foreign currency translations generally produce results opposite to the economic effects of exchange rate changes. After discussing the effects of currency fluctuations on firm value from the perspective of current economic theories, I test whether and how the foreign currency translation adjustment is actually related to changes in value in the manufacturing sector. In doing so, I take the accounting process generating the translation adjustment as a given and analyze the adjustment in view of generally accepted economic theories.

Consistent with existing economic theories, I find that, on average, a positive translation adjustment is associated with a loss of value instead of a creation of value. I also find that the negative association between the translation adjustment and change in value is largely attributable to those firms that are the most labor-intensive. Because such firms are likely to have the most rigid (or "sticky") input factor prices, economic theories predict that they should be the most affected by currency fluctuations. The study raises questions about an accounting income computation that recognizes a positive translation adjustment as an increase in income.

I express my gratitude to Douglas Schroeder (my chair) for his guidance and advice. I also thank Anil Arya, Orie Barron, Anne Beatty, Greg Clinch (the associate editor), Peter Easton, Paul Fisher, Steven Huddart, Bin Ke, James McKeown, Jinhan Pae, Siew Hong Teoh, Richard Young, Andrew Sbaraglia, and two anonymous referees for helpful comments and discussions. This paper is based on my dissertation at Ohio State University.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted November 2001
Accepted March 2003

The remainder of the paper is organized as follows. The next section analyzes the accounting and the economic effects of foreign currency fluctuations on firm value. Section III provides a review of existing studies on the foreign translation adjustment. Section IV describes the research design. The results are reported in Section V. The study concludes in Section IV.

II. ANALYSIS OF THE ACCOUNTING AND THE ECONOMIC EFFECTS OF FOREIGN CURRENCY FLUCTUATIONS

The Accounting Effect of Foreign Currency Fluctuations

In October 1975, the FASB issued SFAS No. 8, *Accounting for the Translation of Foreign Currency Transactions and Foreign Currency Financial Statements*, following the collapse of the fixed exchange rate regime.[1] SFAS No. 8 was severely criticized for treating the foreign exchange gain or loss as a component of net income. In December 1981, the FASB responded by issuing SFAS No. 52, *Accounting for Foreign Currency Translation*, which supersedes SFAS No. 8. The new pronouncement requires the use of the current rate method and the exclusion of translation adjustments from income when a foreign subsidiary uses a foreign currency as its functional currency.[2] The temporal rate method is still required by SFAS No. 52 in the case where the U.S. dollar is deemed the functional currency of the subsidiary and in the case where the subsidiary is located in an hyperinflationary economy. The initial decision as to which currency is the functional currency rests with the parent company.

In general, appreciation (depreciation) of the local currency results in an increase (decrease) in the book value of net assets under the current rate method and a positive (negative) foreign translation adjustment, suggesting an increase (decrease) in firm value.[3] The accounting argument for a positive association between foreign currency fluctuations and firm value focuses on the assets alone, without regard to the role of these assets in the production process. However, the effects of currency fluctuations on the value of a foreign subsidiary's plant and equipment are not clear. When the assets of a foreign subsidiary are imported into the host country or built by U.S. firms, as is usually the case, they are not quoted in the local currencies and, thus, their values in U.S. dollars are not much affected by fluctuations of the local currencies. Even if the assets are bought in the host country, because they are production inputs and not speculative assets, a depreciation (appreciation) of the local currency should, if anything, benefit (hurt) a going concern because the replacement costs of the assets are lower (higher).[4] White et al. (1998, 859) illustrate the effect of the current rate method on the book value of assets as follows:

---

[1] SFAS No. 8 prescribed the temporal rate method. Under this method, monetary items are remeasured at the current exchange rate and nonmonetary items are remeasured at historical rates. Amortization of intangibles, cost of goods sold, deferred charges, and depreciation are remeasured at historical rates. In general, other expenses and revenue are remeasured at the average exchange rate. Because the accounts are remeasured at different rates, in general the basic accounting identity no longer holds. A transaction adjustment gain or loss account is designed to reestablish the identity. Any adjustment gain or loss is posted to net income.

[2] Under the current rate method, all assets and liabilities are translated at the current rate and owners' equity at historical rates. Income statement items are translated at the rate in effect at the time the item is recognized. Because the balance sheet accounts are translated at different rates, in general, the basic accounting equation does not hold. To reestablish the equilibrium, a translation adjustment is made to owners' equity. Note that the translation adjustment results from the translation of a foreign entity's financial statements from the functional currency to U.S. dollar. A transaction gain or loss is recognized for the effect of exchange rate changes on transactions denominated in currencies other than the functional currency.

[3] Ignoring the effect of net capital contribution, the foreign translation adjustment can be expressed as: $ADJ_t = (e_t - e_{t1})NA_{t-1} + (e_t - e_{at})NI_t$, where NA is net assets, NI is net income, e is the exchange rate (U.S. dollar/unit of foreign currency), and e_a is the average exchange rate. Except in the extreme case where earnings is negative and net assets are also negative or less than half the absolute value of earnings, depreciation of the local currency implies an adjustment loss, and vice versa. This is based on the assumption that the average exchange rate, e_a, is computed as half the sum of the beginning and ending exchange rates, e_{t-1} and e_t. In any case, depreciation of the local currency implies an adjustment loss, unless earnings is negative and net assets are also negative or less than the absolute value of earnings.

> Assume that a company builds two identical factories, at identical initial costs, one in the United States and one in Country G, where it has a subsidiary. Assume also that the currency of Country G appreciates relative to the dollar, rising 50% over the next five years. If the currency of Country G is the functional currency for that subsidiary, the "cost" of the factory in that country will be 50% higher than the cost of the U.S. plant.
>
> Presumably, the rise in Country G's currency is due to a lower rate of inflation. But the inflation rate does not change the asset's historical cost. The higher carrying amount of the factory in Country G (and the higher depreciation expense as well) is not logical. It is equally absurd for a factory in a country whose currency depreciates against the dollar to decline steadily in carrying amount. Yet this is the consequence of the application of the all-current rate method.

The Economic Effect of Foreign Currency Fluctuations

From an economic standpoint, fluctuations in exchange rates affect the value of a foreign subsidiary primarily due to the effects of exchange rates on production costs in U.S. dollars. When the local currency depreciates, production costs fall. Similarly, when the local currency appreciates, production costs rise.[5]

Depending on the degree of integration between a foreign subsidiary and its parent, the subsidiary can be classified as either an integrated foreign operation (IFO) or a self-sustained foreign operation (SSFO). An IFO tends to incur production costs in the host country, but generate revenues in the U.S. Hence, when the foreign currency depreciates, an IFO benefits because it becomes cheaper to produce in the host country due to the increase in the relative purchasing power of the U.S. dollar.[6] However, under the current rate method, depreciation of the local currency implies an adjustment loss rather than a gain because net assets are translated at a lower rate. Thus, the postulate that the translation adjustment is a source of value added clearly does not hold for an IFO if management treats it as an independent operation and uses the currency of the country where the subsidiary is located as the functional currency.[7]

---

[4] See Beaver and Wolfson (1984) and Selling and Sorter (1983) for more discussion on the substance of the foreign translation adjustment.

[5] See Capdevielle (1989), Condon et al. (1990), Corbo (1985), Crutchley et al. (1991), Gomez-Ibanez and Harrison (1982), Jenner (1981), Kmitch et al. (1995), Le Fort (1988), Mazumdar (1993), Stolper and Samuelson (1941), and Van Wijnbergen (1985).

[6] Under the purchasing power parity (PPP) theorem, the currency effect should be counterbalanced by the effect of inflation because the rate of inflation in the local economy and the strength of the local currency move in opposite directions (cf. Abuaf and Jorion 1990; Adler and Dumas 1983; Isard 1977; Shapiro 1983). However, the rate of cost inflation generally does not match the rate of depreciation of the local currency because input prices, particularly wages, tend to be sticky or at least to lag overall price increases due to factors such as barriers to the free movement of labor, long-term employment contracts, and government response lags in passing new legislation. The empirical evidence shows that PPP does not hold exactly in any of its forms, whether it is the law of one price, the absolute PPP, or the relative PPP. Discussions and evidence on the relative inertia of nominal wages in foreign markets can be found in Condon et al. (1990), Corbo (1985), Le Fort (1988), Mazumdar (1993), Stolper and Samuelson (1941), and Van Wijnbergen (1985). Absence of rigidity in real wages and other non-internationally traded goods is the primary reason advanced to explain why the pass-through from exchange rate to inflation is not complete. Depending on the degree of market imperfection, particularly labor mobility, induced cross-country production cost differentials may persist for very long periods.

The main issue arises, however, in the case of an SSFO. An SSFO tends to borrow, repay, and sell in foreign currencies. A weaker local currency would result in not only lower production costs, but also lower revenues when translated into U.S. dollars. Thus, at first glance, it may appear that currency fluctuation should not have a material effect on the economic value of an SSFO. The revenue effect should counterbalance the cost effect. However, because input prices are stickier than output prices,[8] appreciation of the local currency negatively affects the value of an SSFO even if the operations of the subsidiary are indeed self-sustained.[9]

Appreciation of the local currency makes it harder to sell in an open market because foreign goods become relatively cheaper (see Aggarwal 1981).[10] To remain competitive, the firm may have to reduce sales price. Labor costs will also be higher; however, as I have explained above, labor costs cannot be adjusted as easily as sales prices. Therefore, appreciation of the local currency is likely to result in decreasing profit margins and greater competition with foreign firms leading to a decline in the value of the firm. Similarly, depreciation of the local currency is likely to result in increasing margins and less competition with foreign firms leading to an increase in the value of the firm. In general, as long as the local markets are not closed, a manufacturing firm is better off operating in a weak foreign currency environment, no matter where it sells its goods. In an internationally competitive market, the world market determines real output prices, while, because of barriers to the free movement of labor, local markets determine real wages.

One objective of SFAS No. 52, as stated by the FASB, is "to provide information that is generally compatible with expected economic effects of a rate change."[11] However, McClenahen and Whenmouth (1988) note that, "The weaker dollar has lowered the prices on U.S. goods shipped to world markets, leading to increased sales for U.S. exporters. U.S. firms manufacturing abroad, however, are forced to pay more dollars to meet foreign currency payrolls, resulting in less profits to be brought back to the U.S." Because the local currencies appreciate, the offshore subsidiaries report positive translation adjustments. However, U.S. manufacturers operating in the countries where the local currencies appreciated are hurt. The following example by White et al. (1998, 861) further illustrates the point:

> Consider ... a British manufacturing subsidiary whose output is sold entirely in Great Britain and whose costs are incurred entirely in pounds sterling. If the value of the pound sterling rises, imports will enter, taking market share from the British subsidiary. The result is likely to be lower sales, lower earnings (price cutting may be necessary to keep market share), and lower cash flows. Once again, with sterling as the functional

---

[7] An IFO is more likely to use the U.S. dollar as its function currency and, therefore to report under the temporal rate method. However, the choice of the functional currency depends largely on management discretion.

[8] cf. Condon et al. (1990), Corbo (1985), Le Fort (1988), Mazumdar (1993), Stolper and Samuelson (1941), and Van Wijnbergen (1985).

[9] Gallardo et al. (2000) find that, over the period 1971 to 1992, the profit margins of Mexico's manufacturing firms increase when the Mexican peso depreciates. An increase in the value of the dollar relative to the Mexican peso is associated with an overall price increase in Mexico. However, Gallardo et al. (2000) explain that Mexican manufacturers benefit because the prices of imported goods rise, relaxing pressure from foreign competition, and sale prices increase faster than labor costs. Clarida (1997) even shows that real currency depreciation should boost the profits of local manufacturers regardless of whether pass-through is complete, partial, or nonexistent.

[10] Recent popular press references include Hagenbaugh (2002), Moffett and Sesit (2002), Samuelson (2002), and Sender (2002).

[11] SFAS No. 52, *Foreign Currency Translation*, FASB (1981, para. 4).

currency, the U.S. dollar net assets will increase. Yet the value of the British subsidiary has almost certainly declined, despite the higher currency exchange rate.

III. PRIOR STUDIES

The extant accounting literature focuses more on the stock market reaction to the foreign exchange adjustment under SFAS No. 52 as opposed to SFAS No. 8. Collins and Salatka (1993) find that the perceived noise created by the inclusion of the currency adjustment in net income negatively affects earnings quality under SFAS No. 8. This conclusion is not supported by Soo and Soo (1994). Soo and Soo (1994) find no difference in the market's valuation of the foreign exchange adjustments generated under the two statements. However, consistent with Collins and Salatka (1993), Bartov (1997) finds that "SFAS 52 has resulted in a significant improvement in the valuation relevance" of earnings. Extant studies examine the market reaction to the foreign exchange adjustments. However, measures of the market's expectation for these adjustments are not readily available. My study complements the unexpected return/earnings literature by using a return/earnings association approach instead. It has the advantage of not being bound by the need to specify unexpected components of the foreign translation adjustment, earnings and returns, and not requiring precise information release dates.

Dhaliwal et al. (1999) use the return/earnings association approach to examine whether adding the foreign translation adjustment to net income increases the association of earnings with return. They also find no evidence that adding the translation adjustment to net income affects the return/earnings association. As Soo and Soo (1994) argue, net income is much larger than the foreign translation adjustment; by adding the translation adjustment to net income, the market effect associated with the adjustment may be overshadowed by fluctuations in net income.[12]

My study also complements the extant literature by formulating and testing a theory of the economics of the foreign translation adjustment. While prior studies implicitly postulate a positive association between change in firm value and the foreign translation adjustment, the analysis of the accounting and economic effects of currency fluctuations leads me to hypothesize an inverse relationship between the foreign translation adjustment and change in firm value. This prediction is based on the premise that depreciation of a local currency entails (1) a negative translation adjustment (the accounting effect) and (2) an increase in the value of a foreign operation (the economic effect). The first condition holds in most cases, except in the extreme instances where earnings is negative and net assets are also negative or considerably smaller than the absolute value of earnings. The second condition only requires that the local economy be open to competition, assuming limited mobility of labor. Note that, to hold, the second premise does not require that the foreign subsidiary export its output back to the U.S.; it holds as long as the local economy is not closed.[13]

IV. RESEARCH DESIGN

Modeling of the Test Variable

To analyze the incremental valuation relevance of the foreign translation adjustment, I use the following return association model:

---

[12] Bodnar and Weintrop (1997) discuss various reasons why different coefficients should be expected on foreign and domestic components of earnings. The case for different coefficients is even stronger for the foreign translation adjustment. The translation adjustment does not reflect gain or loss from operations but instead the effect of fluctuations in exchange rates.

[13] See "Discussion of Related Studies" for further discussions of the extant literature.

$$R_{i,t} = \beta_0 + \beta_1 NI_{i,t} + \beta_2 ADJ_{i,t} + e_{i,t}, \quad (1)$$

where $R_{i,t}$ is the fiscal-year return; $NI_{i,t}$ is the reported net income (Compustat item 18) and $ADJ_{i,t}$ is the translation adjustment (change in Compustat item 230). I scale the right-hand side variables, including the intercept, by beginning market value, yielding a deflated regression of cum-dividend change in value on net income and the translation adjustment.

I am interested in the sign of the coefficient on *ADJ*. Based on the economic rationale that I have formulated earlier, I hypothesize an inverse relationship between the translation adjustment and change in firm value. Hence, I expect β_2 to be negative.

Long-Window Association

The most common alternative to the hypothesis that the foreign translation adjustment represents value added is that the adjustment is a noisy measure that has no value-relevance. For those who maintain that second point of view, any documented association between change in firm value and the translation adjustment would be spurious. The contention is that the market will not pay much attention to short-term fluctuations in exchange rates. This argument suggests that return and the translation adjustment should be uncorrelated; and any documented association between them would be spurious and, therefore, vanish as the holding horizon is extended. To address this concern, I extend the holding horizon from one year to two and five years.[14] I look at how change in firm value and the translation adjustment are correlated over the long run.

Control Variables

Management has some discretion in the initial decision to report under either the temporal or the current rate method. Under the temporal rate method, the effect of the change in exchange rate goes into the transaction gain or loss account instead of the translation adjustment account. The discretion is limited to the year of the initial determination of the functional currency. Hence, I do not expect self-selection to be a serious problem. However, as a precautionary measure, I try to mitigate potential biases by including the transaction gain or loss in the regressions. In addition, I allow for different slopes on the translation adjustment for firms that report under both the current rate method and the temporal rate method, and those that report only under the current method. Then, I test for the extent to which the slope on the translation adjustment is different across the two groups.[15]

The coefficient on the translation adjustment may capture the effect of foreign exposure. In that case, any observed association with return may not necessarily be due to the economic effect that I predict. In addition, the effect of the translation adjustment is likely to be affected by the size of a firm's foreign operations. The inclusion of the transaction gain or loss in the model helps alleviate potential biases due to the association between foreign exposure and the effects of the translation adjustment. To further mitigate the problem, I include foreign income tax in the model. I use foreign income tax as a surrogate for foreign exposure because of the constraints on obtaining hedging data and since foreign income tax data are easily available. In addition, the use of foreign income tax is justified because the theoretical approach of the study concentrates on the effect of foreign currency fluctuations on firms' income. Thus, I extend Regression (1) as follows:[16]

---

[14] Return is compounded over the respective horizon. The explanatory variables are cumulated over the respective horizon. The cumulated earnings numbers should be adjusted for dividend since dividend is factored in the return measurement. But as a practical matter, this adjustment is generally not carried out. One reason is that there may not be a proper way to do it. In addition, Easton et al. (1992) report that adjusting earnings for dividend has no impact on their results.

[15] I study the value relevance of the foreign translation adjustment (current rate method adjustment) and not the remeasurement gains or losses (temporal rate method adjustment). It is possible to formulate a testable hypothesis about the coefficient on the translation adjustment variable because the sign of the adjustment is determined by the direction of exchange rate changes, assuming that net assets are positive. But no such prediction can be made for the remeasurement gain or loss variable because under the temporal rate method the net asset components are remeasured at different rates depending on whether they are classified as monetary or nonmonetary items. In addition, the remeasurement gains or losses are combined in the financial statements with gains or losses from currency transactions such as forward contracts.

$$R_{i,t} = \beta_0 + \beta_1 NI^*_{i,t} + \beta_2 ADJ_{i,t} + \beta_3 FTAX_{i,t} + \beta_4 TADJ_{i,t} + e_{i,t}, \qquad (2)$$

where NI^* is net income (Compustat item 18) minus the transaction gain or loss (Compustat item 150), $TADJ$ is the transaction gain or loss, and $FTAX$ is the foreign income tax (Compustat item 64). All the right-hand side variables, including the intercept, are deflated by beginning market value. This yields a regression of cum-dividend change in firm value on net income, foreign income tax, transaction gains or losses, and translation adjustment, with both sides of the regression scaled by beginning market value.[17] To mitigate the effects of outliers and potential errors in the data, I winsorize the top and bottom 1 percentiles of all the variables.

The use of foreign income tax as a proxy for foreign exposure has two main advantages. First, it is reported on Compustat for most multinational firms. Second, because I am interested in the relevance of the foreign translation adjustment as an income measure, it is useful to use an income-based measure to control for foreign exposure. However, the use of foreign tax has also some limitations. Particularly, cross-sectional differences in tax rates, tax elections, transfer price strategies, and hedging strategies make foreign tax a noisy proxy of foreign exposure. In addition, because the transaction gain or loss includes not only gain or loss from translation under the temporal method, but also gain or loss from currency transactions, that variable is likely to be a noisy measure of foreign exposure and the effect of a firm's choice to report under the temporal method. If the control variables sufficiently capture the intended effects, then even though they are noisy, the coefficient on *ADJ* (the variable of interest) is still unbiased. However, my test results are conservative because the standard errors are upward-biased. In that case, the noise in the control variables works against finding support for my hypothesis. To assess the effect of the noise in the foreign tax and the transaction gain or loss variables, I run the regression with and without these variables.

Another potential problem is that return may be nonlinear in the control variables. To address this issue and other potential specification problems, I submit the results to various diagnostic tests, including accommodation for potential nonlinearity in the control variables, heteroscedasticity, and cross-sectional correlation. Furthermore, if the control variables are noisy, then the coefficient on *ADJ* is biased. To assess the effect of the potential bias, I use a limited sample of firms for which I obtain data on foreign assets to examine the sensitivity of the results to the use of foreign assets as a proxy for foreign exposure.[18]

---

[16] Easton and Harris (1991) find that both earnings and change in earnings has incremental explanatory power for return. However, including different alternative measures for the change in earnings in the regression did not qualitatively affect the results.

[17] My regression of interest is: $P_{i,t} - P_{i,t-1} = \beta_0 + \beta_1 NI^*_{i,t} + \beta_2 ADJ_{i,t} + \beta_3 FTAX_{i,t} + \beta_4 TADJ_{i,t} + e_{i,t}$ with none of the variables deflated. Because of potential scale and heteroscedasticity problems, I deflate both sides of the regression by beginning price ($P_{i,t-1}$) (cf. Christie 1987; Barth and Kallapur 1996; Brown et al. 1998; Easton 1998; Easton and Sommers 2003). Essentially, I estimate the regression of interest by weighted-least-square, using $P_{i,t-1}$ as the deflator. To assess the effect of omitting the standard OLS intercept, I include a column of ones to the deflated model; the results of the paper do not change qualitatively.

Controlling for Unknown Risk Factors

The primary research design and question are formulated in terms of the association between the translation adjustment and change in firm value. I use raw return and avoid framing the study in terms of market reaction to the translation adjustment because, while one can try to estimate unexpected return, there is no reliable way to adjust the translation adjustment for market expectation. However, to address concerns that the translation adjustment may be related to some unknown risk that I have not controlled for, I estimate Regression (2) with abnormal return as the dependent variable.

The abnormal return is defined as the excess return over the median return of a matched portfolio. The matched portfolio consists of firms that have the same two-digit SIC code, similar book-to-market ratio, and similar size. Each year, I form 20 portfolios based on two-digit SIC code using all manufacturing firms that have beginning of the year price per share, number of shares outstanding and book value of equity on Compustat. Next, I divide each of the 20 industry portfolios into two based on whether a firm's book-to-market ratio is above its industry median. Then, I sort each of the 40 book-to-market/industry portfolios into two based on whether a firm's total market capitalization is above the median of its book-to-market/industry portfolio. Finally, abnormal return is computed as the difference between a firm's fiscal year return and the median return of the comparable book-to-market/size/industry-based portfolio.

Sample Selection

The study covers the period 1985 to 2001. The sample includes multinational manufacturing firms (SIC codes: 2000–3999) that have return information on CRSP. I also require that a firm have the same fiscal year-end over the sample period. In addition, a firm must report the cumulative translation adjustment, the lag cumulative translation adjustment, the transaction gain or loss, net income, and the foreign income tax. A multinational firm is defined for the purpose of this study as a firm that has a nonzero cumulative adjustment in a given year. This yields 4,972 observations.

I limit the sample to manufacturing firms because my hypothesis is based on the relative stickiness of input prices, particularly wages. These costs are likely to be more substantial in the manufacturing sector. In addition, manufacturing firms are most likely to be affected by exposure to foreign assets and liabilities. In Table 1, I present the distribution of the sample by industry and year. There is no strong industry clustering. The largest concentrations are in industrial machinery and equipment (21.88 percent of the sample), chemicals and allied products (17.36 percent of the sample), instruments and related products (14.94 percent of the sample), and electrical and electronic equipment (14.56 percent of the sample).

---

[18] I also use foreign earnings as an additional proxy for foreign exposure using a sample of 4,000 firm-years that have observations on foreign earnings (Compustat data item 273) and the other required variables. In Regression (2), I separate net income into domestic and foreign earnings and allow for separate coefficients on the two income components. The results of the paper remain qualitatively the same. The use of foreign earnings as a proxy for foreign exposure is hindered by the frequent incidence of negative earnings. Hence, I also use the absolute value of foreign earnings as a surrogate for foreign exposure. The conclusion of the paper has not changed. The sample includes 612 firm-years that have negative foreign earnings (foreign earnings is the sum of all foreign operations' earnings). I replicate the analyses after removing these firm-years. Again, the results of the paper remain qualitatively the same.

TABLE 1
Sample Distribution

Panel A: Distribution of the Sample by SIC Code

SIC Code	Industry Grouping	# of Observations	% of Observations
2000–2099	Food and kindred products	120	2.414
2200–2299	Textile mill products	46	0.925
2300–2399	Apparel and other textile products	52	1.046
2400–2499	Lumber and wood products	5	0.101
2500–2599	Furniture and fixtures	64	1.287
2600–2699	Paper and allied products	129	2.595
2700–2799	Printing and publishing	69	1.388
2800–2899	Chemicals and allied products	863	17.357
2900–2999	Petroleum and coal products	147	2.957
3000–3099	Rubber and plastics products	172	3.459
3100–3199	Leather and leather products	7	0.141
3200–3299	Stone, clay, glass, and concrete products	73	1.468
3300–3399	Primary metal industries	114	2.293
3400–3499	Fabricated metal products	212	4.264
3500–3599	Industrial machinery and equipment	1,088	21.882
3600–3699	Electrical and electronic equipment	724	14.562
3700–3799	Transportation equipment	240	4.827
3800–3899	Instruments and related products	743	14.944
3900–3999	Miscellaneous manufacturing industries	104	2.092
Total		4,972	100.002[a]

Panel B: Distribution of the Sample by Year

1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
295	295	316	307	288	305	284	297	310	310	315	313	339	279	249	254	216

---

[a] The percentages do not sum to 100 due to rounding.

V. RESULTS

Descriptive Statistics

Descriptive statistics are reported in Table 2. The median (mean) return for the sample is 10.54 (15.70) percent and the median (mean) earnings deflated by beginning price is 0.0593 (0.0327). Both the transaction gain or loss and the translation adjustment have negative means and medians. Their deflated medians (means) are –0.0005 (−0.0011) and −0.0006 (−0.0007), respectively.

Regression Results

Table 3 reports the results of the regressions of return on the translation adjustment. The variables are measured over one-year, two-year, and five-year horizons.[19] For the one-year horizon, I use two models. In Model 1, return is the fiscal-year raw return. In Model 2, return is adjusted for potential risk factors. I find a significantly negative association between the foreign translation adjustment and change in firm value over all three measurement horizons, rejecting both the hypothesis that a positive translation adjustment represents value created and the hypothesis that the translation adjustment has no value relevance. The negative association between the translation adjustment and change in firm value over the long horizon is very much consistent with evidence provided by Baillie and Bollerslev (1989), Darby (1982), Liu and Maddala (1992), Meese and Singleton (1982), Rogalski and Vinso (1977), and Urrutia (1992). These studies show that, though exchange rates are unstable, changes in exchange rates have strong permanent components.

TABLE 2
Descriptive Statistics

	Mean	Std. Dev.	Minimum	Median	Maximum
RETURN	0.1570	0.4693	−0.7736	0.1054	2.8750
NI^*	0.0327	0.1271	−1.0327	0.0593	0.2998
FTAX	0.0134	0.0161	−0.0104	0.0090	0.1101
TADJ	−0.0011	0.0072	−0.0454	−0.0005	0.0387
ADJ	−0.0007	0.0199	−0.1131	−0.0006	0.0999
CADJ	−0.0112	0.0404	−0.3013	−0.0040	0.1401
ABCADJ	0.0241	0.0343	0.0001	0.0119	0.3013
RADJNI	0.2674	0.5702	0.0001	0.0967	5.7049

All the variables, except *RETURN* and *RADJNI*, are deflated by beginning market value. The statistics reported in this table are measured over one year. The mean, median, and standard deviations are computed after winsorizing the top and bottom 1 percentiles of all the variables.

RETURN = 12-month fiscal year cum-dividend return;
NI^* = *NI* minus *TADJ* ;
NI = net income before extraordinary items (Compustat item 18);
TADJ = transaction gain or loss (Compustat item 150);
FTAX = foreign income tax (Compustat item 64);
CADJ = cumulative translation adjustment (change in Compustat item 230);
ADJ = change in *CADJ*;
ABCADJ = absolute value of *CADJ*; and
RADJNI = ratio of the absolute value of *ADJ* to the absolute value of *NI*.

---

[19] I report results using overlapping data for the two-year and five-year horizons. Limiting the sample to non-overlapping data yields qualitatively the same results.

TABLE 3
Conditional Association between the Translation Adjustment and Change in Firm Value

$$R_{i,t} = \beta_0 + \beta_1 NI^{*}_{i,t} + \beta_2 ADJ_{i,t} + \beta_3 FTAX_{i,t} + \beta_4 TADJ_{i,t} + e_{i,t}$$

	Expected Sign	1 Year-Horizon (n = 4,972)				2 Year-Horizon (n = 4,293)		5 Year-Horizon (n = 2,639)	
		Model 1		Model 2					
β_0	NA	2.678 [0.000]	1.909 [0.000]	3.045 [0.000]	2.715 [0.000]	4.145 [0.000]	3.737 [0.000]	7.393 [0.000]	6.132 [0.000]
β_1	+	1.345 (0.000)	1.135 (0.000)	1.188 (0.000)	1.096 (0.000)	1.676 (0.000)	1.407 (0.000)	2.207 (0.000)	1.987 (0.000)
β_2	–	−1.032 (0.001)	−1.002 (0.001)	−0.881 (0.026)	−0.976 (0.016)	−1.423 (0.000)	−1.374 (0.000)	−1.084 (0.013)	−1.402 (0.002)
β_3	+	–	4.056 (0.000)	–	1.791 (0.000)	–	3.272 (0.000)	–	1.446 (0.000)
β_4	NA	–	−1.216 [0.008]	–	0.206 [0.433]	–	−0.721 [0.388]	–	−0.441 [0.573]
Adj. R^2		0.145	0.173	0.079	0.082	0.266	0.297	0.526	0.534

One-tailed p-values are reported in parentheses and two-tailed p-values in brackets.

R is the fiscal-year-return. In Model 2, return is adjusted for potential unknown risk factors. The other variables are defined in Table 2. The horizons indicate the periods over which the variables are measured. The right-hand side variables, including the intercept, are deflated by beginning market value. This yields a regression of cum-dividend change in firm value on net income, foreign income tax, transaction gains or losses, and translation adjustment, with both sides of the regression scaled by beginning market value.

The results are reported in Table 5. The coefficients on the foreign translation adjustment are −3.072, −0.658, and −0.798 for the first, second, and third terciles, respectively. The coefficients are significant for the first and third terciles and insignificant for the second one. F-tests comparing the coefficient for the lower tercile with the coefficients for the second and third terciles are significant at the 0.027 and 0.016 levels, respectively. The very strong coefficient for the firms that have the lowest book-to-market ratios is consistent with the suggestion that the association between the translation adjustment and firm value is stronger for firms whose values are tied up in future earnings.[23]

I provide further insights into whether the negative association between the translation adjustment and change in firm value is due to the opposite effects of exchange rate on the adjustment and future profitability. First, I examine the relevance of the translation adjustment for future foreign earnings. Then, I investigate the robustness of the main results to the inclusion of future foreign earnings as a control variable. I expect the translation adjustment to be negatively correlated with future earnings. Future foreign earnings is in part the economic effect of the change in the value of the local currency. Therefore, by including future foreign earnings in the model, I expect the effect of the future earnings variable to

TABLE 5
Conditional Association between the Foreign Translation Adjustment and Change in Firm Value by Tercile of Lag Book-to-Market Ratio
(n = 4,972)

$$R_{i,t} = \beta_0 + \beta_1 NI^*_{i,t} + \beta_2 ADJ_{i,t} + \beta_3 FTAX_{i,t} + \beta_4 TADJ_{i,t} + e_{i,t}$$

	First Tercile	Second Tercile	Third Tercile
β_0	4.891 [0.000]	2.910 [0.000]	0.990 [0.000]
β_1	1.515 (0.000)	2.350 (0.000)	0.701 (0.000)
β_2	−3.072 (0.003)	−0.658 (0.144)	−0.798 (0.017)
β_3	4.675 (0.000)	0.336 (0.298)	4.787 (0.000)
β_4	−0.822 [0.536]	1.353 [0.452]	−2.548 [0.021]
Adj. R^2	0.129	0.263	0.191

One-tailed p-values are reported in parentheses and two-tailed p-values in brackets.
R is the fiscal-year return. The other variables are defined in Table 2. All the variables are measured in year t and the low/high labor-intensive partition is made in year $t-1$. The right-hand side variables, including the intercepts, are deflated by beginning market value.

---

[23] I also condition the analysis on the level of the firms' foreign asset exposure. I obtain a sample of 1,263 firms for which data on foreign assets are available on the Compustat Segments file. I first estimate the following regression model: $R_{i,t} = \beta_0 + \beta_1 NI_{i,t} + \beta_2 ADJ_{i,t} + \beta_3 FASSET_{i,t} + e_{i,t}$, where *FASSET* is the proportion of assets used in foreign operations (foreign assets divided by total assets). The results show that the coefficient on the translation adjustment is still negative and significant. However, the coefficient on *FASSET* is insignificant. One potential problem is that foreign assets capture only the asset exposure and not the liability exposure. Also, the return variable may be nonlinear in foreign assets. To address the second problem, I indirectly control for foreign assets. I classify the firms into high, medium, and low foreign exposures and estimate a separate regression for each group. The coefficient on the foreign translation adjustment is significantly negative for all three groups of firms. However, it is significantly weaker for those firms that have the lowest foreign asset exposure.

dominate the effect of the foreign translation variable. Barth et al. (1996) use a similar procedure in their study of the value relevance of banks' fair value disclosures.

In Table 6, Panel A, I report the association between the translation adjustment and next-period foreign earnings conditional on current-period foreign earnings. I find a negative association between the translation adjustment and next-period foreign earnings before controlling for foreign assets. However, when I condition on the level of foreign assets, the association is significant only for those firms that have the highest foreign asset exposure. In Panel B, I report the association between the translation adjustment and change in value, conditional on current-period total earnings and next-period foreign earnings. The coefficient on next-period foreign earnings is positive and significant for the firms that have the highest foreign asset exposure and the full sample. This is consistent with Beaver et al. (1980) and Beaver et al. (1987); the market anticipates future earnings. However, after

TABLE 6
[QUERY # 12]

Panel A: Conditional Association between the Foreign Translation Adjustment and Next-Period Foreign Income by the Level of Foreign Asset Exposure (n = 598)

$$FNI_{i,t+1} = \beta_0 + \beta_1 FNI_{i,t} + \beta_2 ADJ_{i,t} + e_{i,t}$$

	Full Sample	Low Foreign Exposure	High Foreign Exposure
β_0	−0.011	−0.134	0.249
	[0.943]	[0.497]	[0.252]
β_1	0.797	1.021	0.657
	(0.000)	(0.000)	(0.000)
β_2	−0.279	0.028	−0.431
	(0.059)	(0.532)	(0.016)
Adj. R^2	0.333	0.337	0.354

Panel B: Association between the Foreign Translation Adjustment and Change in Firm Value Conditional on Next-Period Foreign Income (n = 598)

$$R_{i,t} = \beta_0 + \beta_1 NI_{i,t} + \beta_2 ADJ_{i,t} + \beta_3 FNI_{i,t+1} + e_i$$

	Full Sample	Low Foreign Exposure	High Foreign Exposure	Full Sample	Low Foreign Exposure	High Foreign Exposure
β_0	2.431	2.049	3.009	2.274	2.142	2.094
	[0.194]	[0.459]	[0.232]	[0.224]	[0.441]	[0.410]
β_1	2.084	2.265	1.878	1.906	2.215	1.515
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β_2	−3.340	−2.971	−3.705	−2.051	−2.161	−2.186
	(0.043)	(0.263)	(0.031)	(0.167)	(0.341)	(0.150)
β_3	–	–	–	1.011	0.411	1.533
				(0.065)	(0.367)	(0.023)
Adj. R^2	0.100	0.079	0.121	0.102	0.076	0.130

One-tailed p-values are reported in parentheses and two-tailed p-values in brackets.

FNI is foreign income (Compustat data item 273) and *ADJ* is the foreign translation adjustment. *R* is the fiscal-year return and *NI* is net income before extraordinary items. A firm is considered to have low (high) foreign exposure if its proportion of foreign assets is below (above) the sample median. The right-hand side variables, including the intercept, are deflated by beginning market value.

including next-period foreign earnings in the model, the coefficient on the translation adjustment is no longer significant in either the full sample or the sample of firms that have the highest foreign assets exposure. These results are consistent with the hypothesis that the negative association between the translation adjustment and change in firm value is due to the opposite effects of exchange rate on the adjustment and future profitability.

Diagnostic Tests

Nonlinearity in the Control Variables

The results of the studies are based on OLS estimates. Potential problems include nonlinearity in the control variables. Basu (1997), Freeman and Tse (1992), Hayn (1995), Lyon and Schroeder (1992), and many others cast doubt on the appropriateness of OLS in estimating the return/earnings regression. Given that in Regression (2), the translation adjustment variable (*ADJ*) is the variable of interest, the parametric restrictions imposed on the control variables by OLS may unnecessarily result in inconsistent estimates of the coefficient on *ADJ*.

To check for the robustness of the OLS results, I use the partial linear estimation procedure proposed by Robinson (1988). This two-step process estimates the coefficient on *ADJ* by OLS while relaxing the parametric restrictions on the control variables. In the regression of *R* on *NI*, *FTAX*, *TADJ*, and *ADJ*, the first step consists of nonparametrically regressing both *R* and *ADJ* on *NI*, *FTAX*, and *TADJ*.[24] The residuals from the two regressions are obtained, $R\text{-}E(R|NI, FTAX, TADJ)$ and $ADJ\text{-}E(ADJ|NI, FTAX, TADJ)$. A large (small) residual means that the dependent variable (*R* or *ADJ*) is large (small) after controlling for net income, the foreign income tax and the transaction gain or loss. The second step consists of estimating a no-intercept OLS regression of the conditional mean differenced return on the conditional mean differenced translation adjustment. Robinson (1988) shows that such a procedure results in consistent and asymptotically normal coefficient estimates on the test variable, in this case, *ADJ*. Nontabulated results from the semiparametric regressions are consistent with the prediction of the study. The coefficients on *ADJ* are negative and significant over all three measurement horizons.

Cross-Sectional Heteroscedasticity and Cross-Sectional Correlation

Other potential problems with the regression estimates include cross-sectional heteroscedasticity and cross-sectional correlation. Fama and MacBeth (1973) suggest making inferences on the basis of the average coefficients and the standard deviations of annual cross-sectional regressions. One advantage of this procedure is that the variation in the time-series coefficients reflects the cross-sectional correlations in the error terms. However, because the time-series covered by my sample is finite (17 years), the ability to make valid inferences using this procedure is limited. The number of time-series coefficients is too small to make the normality assumption, a necessary condition for making statistical inferences. Nevertheless, in Table 7, I report the coefficients on the foreign translation adjustment from the year-by-year cross-sectional regressions. They are consistent with the other results of the study. They also shed some light on results of prior studies, which I discuss in the next section.[25]

---

[24] The nonparametric regression procedure used follows Hardle (1990). The bandwidth is chosen by generalized cross-validation (Craven and Wahba 1979) to minimize the mean square errors of the nonparametric regression estimate. Refer to Lyon and Schroeder (1992) for a detailed discussion on the nonparametric regression procedure.

[25] I also estimate time-series regressions for each of the 27 firms that have complete observations from 1985 to 2001. The average and median coefficient estimates on the translation adjustment are negative and significant at conventional levels.

TABLE 7
Conditional Association between the Translation Adjustment and Change in Form Value by Year

Model 1: $R_{i,t} = \beta_0 + \beta_1 NI^*_{i,t} + \beta_2 ADJ_{i,t} + e_{i,t}$
Model 2: $R_{i,t} = \beta_0 + \beta_1 NI^*_{i,t} + \beta_2 ADJ_{i,t} + \beta_3 FTAX_{i,t} + \beta_4 TADJ_{i,t} + e_{i,t}$

	Model 1	Model 2
1985	5.437	2.829
1986	3.701	2.693
1987	−0.308	−1.595
1988	−0.666	0.853
1989	−1.184	−0.211
1990	−4.396	−4.364
1991	−2.815	0.489
1992	−1.056	−0.170
1993	−3.962	−1.567
1994	0.046	−0.919
1995	3.002	1.179
1996	−2.786	−0.600
1997	−3.471	−0.797
1998	−1.050	−1.397
1999	−5.009	−3.679
2000	−1.301	2.109
2001	−4.687	−3.299
Mean	−1.206	−0.497
Median	−1.184	−0.600

This R is the fiscal-year-return. The other variables are defined in Table 2. The right-hand side variables are deflated by beginning market value.

An alternative procedure to assess the effects of cross-correlation is the bootstrap approach suggested by Freedman and Peters (1984). A prior application of this bootstrap procedure in the accounting literature can be found in Kross et al. (1990). The bootstrap procedure applies "to econometric models, where the technical difficulties include simultaneity, correlated errors, heteroscedasticity, and dynamics." Nontabulated results show that the mean (median) of the bootstrap coefficient estimates are negative for all three measurement horizons considered in the study. The OLS coefficient estimates are in general slightly higher than the bootstrapped estimates. At the same time, the bootstrapped standard errors are higher than the OLS standard errors. However, t-statistics computed on the basis of the bootstrapped estimates are still significant.[26]

Discussion of Related Studies

Results reported in Table 7 show that, though the regression coefficients on the translation adjustment are negative for most of the years, they are large and positive for the first two years of the sample. This is potentially one explanation for the positive association between abnormal return and the translation adjustment documented by Bartov (1997).

---

[26] I also adjust the standard errors according to White (1980). The use of heteroscedasticity-consistent standard errors does not change the conclusion of the paper.

Bartov's sample period ends in 1990. For the sake of comparison, I estimate my model over the period 1985–1990. The coefficient on the translation adjustment is negative and significant with a one-tailed p-value of 0.028. Apart from some specification differences between Bartov (1997) and the current study, there remains a very important difference between the two studies. Bartov's (1997) sample includes all firms irrespective of their industrial operations. His sample for the SFAS No. 52 regime includes 2,979 manufacturing firms and 1,080 nonmanufacturing firms, whereas my sample is limited to manufacturing firms since I have no prediction for the other industrial sectors. In addition, since there is not an appropriate proxy for unexpected translation adjustment, return and not abnormal return is the appropriate dependent variable.

Soo and Soo (1994) also find a positive association between the market's reaction at the earnings announcement and the foreign translation adjustment. Obviously, there is no reason to expect such an association (particularly, a positive one) since the foreign translation adjustment is not known at the earnings announcement date (cf. Soo and Soo 1994, footnote 5). Even if it were known, there should not be any association because the translation adjustment is the result of currency fluctuations. As exchange rates change, investors should adjust stock prices accordingly; there is no reason why they should wait for the earnings announcement or the report filing date to do so (see Bartov 1997, 632). This is likely one reason why Soo and Soo (1994) find no association between the adjustment and abnormal return around the annual report filing date when the translation adjustment is supposedly known to the market. (See Soo and Soo 1994, footnote 10.)

VI. CONCLUDING REMARKS

The study analyzes whether and how the foreign translation adjustment is related to value for firms in the manufacturing sector. I find that the translation adjustment is value relevant, but not in the direction of the adjustment. A positive adjustment is associated with a loss of value. The study raises concerns about an accounting income computation that recognizes a positive translation adjustment as an increase in income.

Recognizing the translation adjustment as a component of income bears some resemblance to the current treatment of R&D and advertising expenditures. R&D and advertising are recognized as expenses even when they may be value-adding. However, it is important to notice that R&D and advertising involve real expenditures, while the foreign translation adjustment does not. The translation adjustment is added to book value simply for the sake of the double entry system, not because of any need to record a transaction.

The need to articulate the income statement and the balance sheet is one of the reasons suggested for adding the translation adjustment to net income (cf. AAA FASC 1997). The articulation between the two statements is a legitimate concern. However, the argument that adding the translation adjustment to net income is necessary to maintain clean surplus and harmonize the financial statements is not entirely correct. If the translation adjustment does not represent value creation in the firm, then there is no compelling reason to add it to book value in the first place. Note again that the translation adjustment does not involve any transaction but only the need to balance the basic accounting equation. Note also that the need for the adjustment is due mainly to the translation of net assets at current exchange rates (a departure from the historical cost principle). To maintain the equilibrium of the basic accounting equation, another option would be to make the adjustment on the asset-side of the equation. In terms of balancing the basic accounting equation and articulating the income statement with the balance sheet, subtracting the adjustment from assets or adding it to both equity and net income yields the same results. However, since the translation adjustment is related to a decrease in value, subtracting it from assets is more consistent with the economic reality.

REFERENCES

Abuaf, N., and P. Jorion. 1990. Purchasing power parity in the long run. *Journal of Finance* 45: 157–174.

Adler, M., and B. Dumas. 1983. International portfolio choice and corporation finance: A synthesis. *Journal of Finance* 38: 925–984.

Aggarwal, R. 1981. Exchange rates and stock prices: A study of the U.S. capital markets under floating exchange rates. *Akron Business and Economic Review* 12: 7–12.

American Accounting Association Financial Accounting Standards Committee (AAA FASC). 1997. An issue paper on comprehensive income. *Accounting Horizons* 11: 120–126.

Baillie, R., and T. Bollerslev. 1989. Common stochastic trend in a system of exchange rates. *Journal of Finance* 44: 167–181.

Barth, M., W. Beaver, and W. Landsman. 1996. Value-relevance of banks' fair value disclosures under SFAS No. 107. *The Accounting Review* 71: 513–537.

———, and S. Kallapur. 1996. The effects of cross-sectional scale differences on regression results in empirical accounting research. *Contemporary Accounting Research* 13: 527–67.

Bartov, E. 1997. Foreign currency exposure of multinational firms: Accounting measures and market valuation. *Contemporary Accounting Research* 14: 623–652.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24: 3–37.

Beaver, W., R. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting and Economics* 2: 3–28.

———, and M. Wolfson. 1984. Foreign currency translation gains and losses: What effect do they have and what do they mean? *Financial Analysts Journal* 40: 28–36.

———, R. Lambert, and S. Ryan. 1987. The information content of security prices: A second look. *Journal of Accounting and Economics* 9: 139–157.

Bodnar, G., and J. Weintrop. 1997. The valuation of the foreign income of U.S. multinational firms: A growth opportunities perspective. *Journal of Accounting and Economics* 24: 69–97.

Brown, S., K. Lo, and T. Lys. 1999. Use of R^2 in accounting research: Measuring changes in value relevance over the last four decades. *Journal of Accounting and Economics* 28: 83–115.

Capdevielle, P. 1989. International comparisons of hourly compensation costs. *Monthly Labor Review* 112: 10–12.

Christie, A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting and Economics* 16: 231–258.

Clarida, R. 1997. The real exchange rate, exports, and manufacturing profits: A theoretical framework with some empirical support. *International Journal of Finance and Economics* 2: 177–188.

Collins, D., and W. Salatka. 1993. Noisy accounting earnings signals and earnings response coefficients: The case of foreign currency accounting. *Contemporary Accounting Research* 10: 119–159.

Condon, T, V. Corbo, and J. De Melo. 1990. Exchange rate based disinflation, wage rigidity, and capital inflows: Tradeoffs for Chile, 1977–1981. *Journal of Development Economics* 23: 113–131.

Corbo, V. 1985. International prices, wages and inflation in an open economy: A Chilean model. *Review of Economics and Statistics* 67: 564–573.

Craven, P., and G. Wahba. 1979. Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by method of generalized cross-validation. *Numerische Mathematik* 31: 377–403.

Crutchley, C., E. Guo, and R. Hansen. 1991. Stockholder benefits from Japanese-U.S. joint ventures. *Financial Management* 20: 22–30.

Darby, M. 1982. Does purchasing power parity work? *Proceedings of Fifth West Coast Academic/Federal Reserve Economic Research Seminar.* San Francisco, CA: Federal Reserve Bank of San Francisco.

Dhaliwal, D., K. Subramanyam, and R. Trezevant. 1999. Is comprehensive income superior to net income as a measure of firm performance? *Journal of Accounting and Economics* 26: 43–67.

Easton, P., and T. Harris. 1991. Earnings as an explanatory variable for return. *Journal of Accounting Research* 29: 19–36.

———, ———, and J. Ohlson. 1992. Aggregate accounting earnings can explain most of security return. *Journal of Accounting and Economics* 15: 119–142.

———. 1998. Discussion of: Revalued financial, tangible, and intangible assets: Association with share prices and non market-based value estimates. *Journal of Accounting Research* 36: 235–247.

———, and G. Sommers. 2003. Scale and scale effects in market-based accounting research. *Journal of Business Finance and Accounting* 40: 25–56.

Fama, E., and J. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81: 607–636.

Financial Accounting Standards Board (FASB). 1981. *Foreign Currency Translation.* Statement No. 52. Stamford, CT: FASB.

Freedman, D., and S. Peters. 1984. Bootstrapping a regression equation: Some empirical results. *Journal of the American Statistical Association* 79: 97–106.

Freeman, R., and S. Tse. 1992. A nonlinear model of security responses to accounting earnings. *Journal of Accounting Research* 30: 185–209.

Gallardo, J., M. Anyul, and J. Hernansez. 2000. Profit margins in Mexico's manufacturing industry: An econometric study. *Metroeconomica* 51: 102–121.

Gomez-Ibanez, J., and D. Harrison. 1982. Imports and the future of the U.S. automobile industry. *The American Economic Review* 72: 319–323.

Hagenbaugh, B. 2002. Dollar falls against major currencies. *USA Today* (December 17): B3.

Hardle, W. 1990. *Applied Nonparametric Regression.* Cambridge, MA: Cambridge University Press.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20: 125–153.

Isard, P. 1977. How far can we push the law of one price? *American Economic Review* 67: 942–948.

Jenner, S. 1981. Inflation, exchange rates and the location of automobile manufacturing. *Business Economics* 16: 26–29.

Kmitch, J., P. Laboy, and S. Damme. 1995. International comparisons of manufacturing compensation. *Monthly Labor Review* 118: 3–9.

Kross, W., B. Ro, and D. Schroeder. 1990. Earnings expectations: The analysts' information advantage. *The Accounting Review* 65: 461–476.

Le Fort, G. 1988. The relative price of nontraded goods, absorption and exchange rate policy in Chile, 1974–1982. *Staff Papers, International Monetary Fund* 35: 336–370.

Liu, P., and G. Maddala. 1992. Rationality of survey data and tests for market efficiency in foreign exchange markets. *Journal of International Money and Finance* 16: 366–381.

Lyon , J., and S. Schroeder. 1992. Firm growth and the valuation relevance of earnings levels, earnings innovations, and dividends. *Journal of Accounting, Auditing, and Finance* 7: 531–551.

Mazumdar, D. 1993. Labor market and adjustment in open Asia economies: The Republic of Korea and Malaysia. *World Bank Economic Review* 7: 349–380.

McClenahen, J., and E. Whenmouth. 1988. International: The falling dollar two views. *Industry Week* 236 (January 18): 41–43.

Meese, R., and K. Singleton. 1982. On the unit roots and the empirical modeling of exchange rates. *Journal of Finance* 37: 1029–1034.

Moffett, S., and R. Sesit. 2002. A global journal report: Japan sees hope in weak yen—Export-focused policy could clobber U.S. manufacturers. *Wall Street Journal* (December 9): A14.

Robinson, P. 1988. Root-n-consistent semiparametric regression. *Econometrica* 56: 931–954.

Rogalski, R., and J. Vinso. 1977. Price level variations as predictors of flexible exchange rates. *Journal of International Business Studies* 8: 71–81.

Samuelson, R. 2002. Deflation out of China? *Washington Post* (December 4): A23.

Selling, T., and G. Sorter. 1983. FASB Statement No. 52 and its implications for financial statement analysis. *Financial Analysts Journal* 39: 64–69.

Sender, H., 2002. China resists calls to let yuan float in strong economy. *Wall Street Journal* (December 24): B6.

Shapiro, A. 1983. What does purchasing power parity means? *Journal of International Money and Finance* 2: 295–318.

Soo, B., and L. Soo. 1994. Accounting for the multinational firm: Is the translation process valued by the stock market? *The Accounting Review* 69: 617–637.

Stolper, W., and P. Samuelson. 1941. Protection and real wages. *Review of Economic Studies* 9: 58–72.

Urrutia, J. 1992. Variance ratio tests of random walk for foreign exchange rates. *Economic Letters* 38: 457–465.

Van Wijnbergen, S. 1985. Macro-economic effects of changes in bank interest rates: Simulation results for South Korea. *Journal of Development Economics* 18: 541–554.

White, G., A. Sondi, and D. Fried. 1998. *The Analysis and Use of Financial Statements.* 2nd edition. New York, NY: John Wiley & Sons, Inc.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test of heteroskedasticity. *Econometrica* 48: 817–838.

THE ACCOUNTING REVIEW
Vol. 78, No. 4
2003
pp. 1049–1067

The Implications of Using Stock-Split Adjusted I/B/E/S Data in Empirical Research

Jeff L. Payne
Wayne B. Thomas
University of Oklahoma

ABSTRACT: The purpose of this study is to highlight issues of interest to researchers employing the I/B/E/S earnings and forecast data. I/B/E/S has traditionally provided per share data on a split-adjusted basis, rounded to the nearest penny. In doing so, per share amounts are comparable over time. However, because not all prior forecasts and earnings per share amounts divide precisely to a penny, adjusting for stock splits and rounding to the nearest penny can cause a loss of information. Researchers are prohibited in many cases from determining the amounts actually reported in prior years, leading to misclassified observations. We obtain actual (unadjusted) earnings and forecast data from I/B/E/S and compare results to those generated using the adjusted I/B/E/S data. We replicate prior studies and find that conclusions are affected when using the actual I/B/E/S data.

Keywords: *analysts' forecast errors; earnings; earnings changes; earnings announcement returns; dispersion; misclassification.*

Data Availability: *All data are from public sources.*

I. INTRODUCTION

The purpose of this research is to investigate how using stock-split adjusted analysts' forecasts data can influence inferences drawn in extant research. A large body of research employs analyst forecast data to investigate various financial accounting issues. Financial analysts' forecast data provide observable measures of expected earnings by sophisticated users of financial accounting information (Brown et al. 1987; Brown 1993). As such, researchers use them to examine issues related to the valuation of a firm's securities, management's incentives to manipulate earnings, and the overall quality of a firm's information environment. In order to specify appropriate tests and draw appropriate inferences researchers must understand the conventions used in creating the forecast data. The

---

We are especially thankful to Bob Lipe and two anonymous reviewers whose comments were instrumental in developing the paper. We are also thankful for comments received from Jeff Abarbanell, Fran Ayres, Don Herrmann, Marlys Lipe, and G. Lee Willinger. We gratefully acknowledge the contribution of I/B/E/S Inc. for providing earnings per share forecast data, available through the Institutional Brokers Estimate System. I/B/E/S provides this data as part of a broad academic program to encourage earnings expectations research.

Editor's note: This was accepted by Terry Shevlin, Senior Editor.

Submitted October 2002
Accepted March 2003

procedures employed by data providers might impact research results (Abarbanell and Lehavy 2001).

One of the primary providers of analyst forecast data is Institutional Brokers Estimate System (I/B/E/S).[1] I/B/E/S issues annual (from 1976) and quarterly (from 1984) forecasts. Traditionally, I/B/E/S has provided forecast data on an adjusted basis, rounded to two decimal places.[2] The adjustment controls for the effects of stock splits on per share amounts over time. Rounding to two decimals is presumably done so that all amounts are reported to the nearest penny. While it is important to be aware of both issues (split adjustment and subsequent rounding), the latter issue is potentially more troublesome to researchers. When reported amounts are divided by the split factor, forecast errors that were originally large can appear quite small. Further, because of the rounding, the researcher can recreate the actual reported amounts only if the adjusted amounts have no more than two decimal places. Otherwise, information is lost. For actual amounts rounded to zero by the stock-split adjustment, multiplying by the split factor or scaling by some per share amount will not correct the rounding effect. The problem is best illustrated with the following examples.

Suppose in Year 1 that Company A and Company B have the follow earnings per share and forecast per share amounts:

Example 1:

		Year 1	After a 4-for-1 Stock Split in Year 2* Year 1 (adjusted)
Company A	Earnings	.99	.25
	Forecast	1.00	.25
	Forecast Error	−.01	.00
Company B	Earnings	1.01	.25
	Forecast	1.00	.25
	Forecast Error	.01	.00

* Amounts are rounded to two decimals, as done by I/B/E/S.

In Year 2, both companies have a 4-for-1 stock split. On an adjusted basis, I/B/E/S reports actual earnings per share in Year 1 for Company A and Company B as .25. These numbers are the actual earnings per share divided by four and then rounded to the nearest penny. The adjusted forecast amount for each firm would be reported by I/B/E/S as 0.25. The researcher would then conclude that the forecast errors for both firms are zero.[3,4] The adjustment issue becomes more pronounced as the split factor increases. Suppose Company A and Company B have the following earnings per share and forecast per share amounts in Year 1:

---

[1] Over 1,500 researchers at over 400 institutions worldwide rely on I/B/E/S's analyst forecast data for research (I/B/E/S 2000). The most recent issue of the I/B/E/S *Research Bibliography* contains 575 abstracts from published and ongoing research that use the I/B/E/S database. We reviewed publications for the past four years (1999–2002) in *The Journal of Accounting Research*; *Journal of Accounting and Economics*; *The Accounting Review*; *Journal of Accounting, Auditing and Finance*; and *Review of Accounting Studies*. We found at least 40 articles that used the I/B/E/S summary tapes or First Call in the past four years.

[2] The data provided by I/B/E/S on the Summary files are rounded to two decimals. The data provided on the Detail files are rounded to four decimals, indicating that the rounding issues discussed in this paper are less severe if the Detail files are used. The primary analysts' forecasts data provided by First Call are adjusted for stock-splits and rounded to two decimal places.

Example 2:

		Year 1	After a 64-for-1 Stock Split in Year N* Year 1 (adjusted)
Company A	Earnings	.33	.01
	Forecast	.64	.01
	Forecast Error	−.31	.00
Company B	Earnings	.95	.01
	Forecast	.64	.01
	Forecast Error	.31	.00

* Amounts are rounded to two decimals, as done by I/B/E/S.

By Year N, each firm has an adjustment factor of 64 for year 1 (e.g., six 2-for-1 stock splits since Year 1). For Company A and Company B, I/B/E/S would report zero forecast error in Year 1. In reality, Company A missed its forecast by almost 50 percent and Company B exceeded its forecast by nearly 50 percent. Using the adjusted I/B/E/S data, the researcher cannot distinguish these observations.

When using the adjusted, rounded data provided by I/B/E/S, a researcher could misclassify prior years' analysts' forecast errors. Because stock splits are not random events, the misclassification can bias the empirical results. The purpose of our study is to first provide general evidence about the potential influence of rounding to two decimal places. The results indicate that stock split factors are of sufficient magnitude to affect reported adjusted amounts and the differences between adjusted and actual forecast data are systematically related to firm characteristics such as size and price-to-book ratios. Next, we use I/B/E/S actual data to replicate prior research that used the adjusted data. We find that many prior conclusions are affected by the rounding procedure. Furthermore, other studies with the potential to contribute to the accounting and finance literature may have been abandoned because the adjusted I/B/E/S data led to insignificant results when, in fact, a significant relation was present.

The paper proceeds as follows. In Section II we discuss the research design and report the results of our replications using both the adjusted and actual I/B/E/S forecast data. We conclude the paper in Section III.

---

[3] We use a 4-for-1 split to clearly illustrate the effect of stock splits on adjusted data. Smaller factors could lead to rounding errors as well. For the example shown, a stock split factor as low as 1.07 would cause the negative one cent forecast error to be reported as zero forecast error.

[4] The split adjustment and subsequent rounding can also affect intervals other than zero. For example, when I/B/E/S reports an adjusted forecast error of $.01 for a firm that has a split factor of 4, the actual forecast error might have been $.02, $.03, $.04, or $.05. All of these actual forecast errors would round to $.01 after being divided by 4. There is no way for the researcher to determine using the I/B/E/S data which amount was the actual forecast error. Scaling by price per share reduces, but does not always eliminate, this rounding effect.

II. EMPIRICAL ANALYSIS

The potential for the rounding procedure to cause problems depends on the study's sample characteristics and research design. As shown in the introduction, the rounding procedure employed by I/B/E/S is potentially problematic only for firms that report stock splits or stock dividends (hereafter both denoted as stock splits), otherwise there would be no subsequent rounding of previous amounts. In addition, the problem is exacerbated at zero forecast error, since the researcher cannot "unadjust" the data.[5] Therefore, conclusions are more likely to be affected in research settings that focus on zero forecast error, that involve firms that are more likely to have splits (e.g., larger firms, growth firms, well-performing firms, etc.), and/or that fail to consider the adjustment made by I/B/E/S. In the following sections, we provide general and specific evidence concerning the effects of the rounding procedure by replicating prior published research. For each replicated study, we follow as closely as possible the research design, sampling procedure, variable measurement, statistical tests, and controls for outliers. We determine if inferences change according to whether the adjusted or actual data are employed.

We obtain analysts' forecasts data from I/B/E/S's for this study.[6] One dataset is stock-split adjusted and rounded to two decimals. These are the data provided to academics under I/B/E/S's standard subscription agreement. We also obtain data from I/B/E/S for the same time period, via a special request, that is not stock-split-adjusted. These data represent the analysts' forecasts, earnings information, stock prices, etc., as they were originally entered into the I/B/E/S database. These amounts are not adjusted for subsequent stock splits.[7] The full sample includes 173,286 firm/quarter observations for the years 1984–1999.[8]

General Evidence

We first provide information on the distribution of stock split factors to demonstrate how sample observations are affected by stock splits. As shown in Table 1, the mean factor declines from a little more than 3 in 1984 to just above 1 in 1999.[9] In 1984 the top quartile of firms had split factors in excess of 3, with the top decile in excess of 6.75. These values gradually decline, as expected, but still retain values greater than 1 until 1998 for the upper quartile and stay greater than 1 for the top decile of firms. Notice that the maximum split factor gets as large as 288 in 1990.[10] This does not by itself indicate that the analysts' forecast errors for stock split firms contain rounding errors. The influence of the stock split factor on rounding is determined by several factors including the level of reported earnings, analysts' forecasts, and the related analysts' forecast error. However, the results in Table 1 indicate that stock split factors are of sufficient magnitude to affect reported adjusted amounts as shown in Examples 1 and 2.

---

[5] In addition, when the researcher is interested in measuring the actual forecast error per share (unscaled), it is impossible to determine that amount for many of the observations. Multiplying by the split factor can only approximate the actual forecast error (see footnote 4).

[6] The actual and adjusted per share I/B/E/S data used in our study come from the Summary files.

[7] As a sensitivity test, we replicate the adjusted data by applying the adjustment factor from the adjusted tapes to the actual data. For 90 percent of our sample, our replication of the adjusted data and the adjusted data provided by I/B/E/S agree. Excluding the remaining 10 percent of the observations where we could not replicate the adjusted I/B/E/S data does not affect any of our conclusions. Of the observations that could not be replicated, approximately 90 percent (i.e., 9 percent of the total sample) differs by one cent.

[8] As discussed below, one of the studies we choose to replicate employs annual forecast data over the 1980–1991 period. Following that study's sampling procedures, we also employ annual forecast data since 1980.

[9] We use a version of I/B/E/S data created on March 2001 for our analysis. Therefore, stocks that had stock splits subsequent to 1999 will report a split factor greater than 1 in 1999.

[10] The stock split factor is not monotonically increasing as newer companies such as Dell and Microsoft, who have had numerous stock splits, were added to the I/B/E/S database after 1984.

TABLE 1
Distribution of Stock Split Factors as Reported by I/B/E/S

Year	Mean	Min.	Lowest Decile	Q1	Median	Q3	Highest Decile	Max.
1984	3.041	0.005	1	1	1.3	3	6.75	64
1985	3.093	0.005	1	1	1.5	3	6	160
1986	2.824	0.005	1	1	1.5	3	6	108
1987	2.715	0.010	1	1	1.15	2.44	5	160
1988	2.623	0.010	1	1	1.05	2.25	4.4	162
1989	2.521	0.010	1	1	1.03	2.25	4	96
1990	2.512	0.013	1	1	1.1	2.2	4	288
1991	2.456	0.025	1	1	1.1	2	4	288
1992	2.213	0.025	1	1	1	2	4	144
1993	1.940	0.001	1	1	1	2	3.4	128
1994	1.761	0.001	1	1	1	2	3	128
1995	1.666	0.001	1	1	1	2	3	64
1996	1.499	0.001	1	1	1	1.5	2.2	64
1997	1.372	0.001	1	1	1	1.25	2	32
1998	1.251	0.001	1	1	1	1	2	20
1999	1.169	0.001	1	1	1	1	1.5	20
All	1.939	0.001	1	1	1	2	3.71	288

The stock split factors reported are those from the March 2001 I/B/E/S Summary (adjusted) tape. I/B/E/S defines this variable as the "cumulative factor that has been applied to historical data to adjust for splits and capitalization changes."

Next, we provide general evidence that use of the adjusted data versus actual I/B/E/S data can affect inferences. We focus this analysis on settings where reported analysts' forecast errors equal zero as this represents the adjusted data that is most likely affected by the stock split adjustment. Forecast error is computed as earnings per share for the quarter as reported by I/B/E/S minus the mean forecast of earnings per share as of the final month of the quarter.

Based on our previous discussion, we expect the following differences in the actual and adjusted data:

(1) the adjusted data will overstate the percentage of observations with amounts equal to zero,
(2) the overstatement of observations with amounts equal to zero will be greater in earlier years (because split factors are applied retroactively),
(3) the distributions of the adjusted data will be tighter, and
(4) firms that are more likely to have stock splits (e.g., larger firms and higher growth firms) will have more observations misclassified at zero.

As anticipated, the percentage of observations equal to zero is greater for the adjusted data than the actual data (see Panel A of Table 2). Examination of the first three columns shows that the proportion of zero-adjusted forecast errors (*AdjFE*) is significantly greater than the proportion that actually occurred (*ActFE*), and these differences are greater in earlier years because of the retroactive application of the split adjustment. The statistics reported in the Ratio column, calculated as (*AdjFE*/*ActFE*) − 1, indicate the degree of misclassification. This ratio begins at 165.59 percent in 1984 and gradually declines to 7.40 percent in 1999, with an overall misclassification ratio of 41.55 percent. This pattern is

TABLE 2
Comparisons of the Adjusted and Actual I/B/E/S Data
(n = 173,286)

Panel A:

Year	Percentage of Observations with Zero Forecast Error as Reported by I/B/E/S			Interquartile Range for All Observations as Reported by I/B/E/S	
	AdjFE[a]	ActFE[b]	Ratio[c]	AdjFE	ActFE
1984	8.81%	3.32%[d]	165.59%	−0.11 to 0.04	−0.16 to 0.09
1985	10.28	4.50[d]	128.24	−0.11 0.02	−0.17 0.04
1986	11.81	5.17[d]	128.50	−0.09 0.02	−0.14 0.03
1987	12.31	5.83[d]	111.19	−0.07 0.02	−0.11 0.04
1988	13.07	6.27[d]	108.57	−0.06 0.03	−0.08 0.05
1989	12.74	6.45[d]	97.58	−0.07 0.02	−0.10 0.03
1990	14.65	7.84[d]	86.93	−0.07 0.01	−0.10 0.02
1991	14.72	8.23[d]	78.83	−0.06 0.01	−0.08 0.03
1992	16.06	9.86[d]	62.84	−0.04 0.02	−0.05 0.03
1993	15.37	10.27[d]	49.70	−0.04 0.02	−0.05 0.03
1994	15.95	11.63[d]	37.19	−0.03 0.02	−0.04 0.03
1995	16.39	12.10[d]	35.49	−0.03 0.02	−0.04 0.03
1996	17.02	13.37[d]	27.26	−0.03 0.02	−0.03 0.03
1997	17.70	14.72[d]	20.23	−0.02 0.02	−0.02 0.03
1998	16.73	14.85[d]	12.69	−0.03 0.02	−0.03 0.02
1999	16.11	15.00[d]	7.40	−0.02 0.03	−0.03 0.03
All	15.19%	10.73%[d]	41.55%	−0.04 0.02	−0.06 0.03
Trend[e]	.505	.821[f]			
(p-value)	(.01)	(.01)			

Panel B: Percentage of Observations with Zero Forecast Error as Reported by I/B/E/S

	Market Value (Small/Large)				Price-to-Book (Low/High)			
	AdjFE		ActFE		AdjFE		ActFE	
Year	Small	Large	Small	Large	Low	High	Low	High
1984	7.27%	11.51%[g]	3.48%[d]	2.37%[d]	3.97%	17.43%[g]	1.43%[d]	6.34%[d,g]
1985	9.08	14.11[g]	4.38[d]	4.38[d]	4.89	20.48[g]	2.80[d]	8.46[d,g]
1986	8.72	16.16[g]	4.71[d]	4.86[d]	5.25	22.08[g]	2.91[d]	8.26[d,g]
1987	10.22	15.77[g]	5.55[d]	5.71[d]	5.93	22.50[d]	3.41[d]	8.71[d,g]
1988	10.99	14.45[g]	5.95[d]	5.49[d]	6.88	22.09[g]	3.36[d]	10.04[d,g]
1989	10.21	15.34[g]	6.15[d]	5.92[d]	5.75	21.96[g]	3.14[d]	10.12[d,g]
1990	12.56	17.76[g]	8.00[d]	7.50[d]	8.12	24.50[g]	4.87[d]	12.39[d,g]
1991	13.36	18.58[g]	7.96[d]	9.16[d,g]	7.98	25.74[g]	4.76[d]	14.68[d,g]
1992	14.63	19.16[g]	10.47[d]	9.07[d,g]	8.83	27.45[g]	5.59[d]	16.24[d,g]
1993	13.60	18.75[g]	10.69[d]	10.36[d]	10.49	24.31[g]	7.89[d]	14.36[d,g]
1994	14.72	18.33[g]	11.90[d]	11.79[d]	12.01	21.84[g]	9.38[d]	14.51[d,g]
1995	14.96	18.42[g]	12.31[d]	11.78[d]	11.40	22.35[g]	9.59[d]	14.84[d,g]
1996	14.59	20.33[g]	12.25[d]	14.23[d,g]	11.67	22.78[g]	9.85[d]	16.76[d,g]
1997	14.77	20.97[g]	13.29[d]	16.07[d,g]	13.31	22.06[g]	12.13	16.71[d,g]
1998	14.33	19.55[g]	13.61	16.26[d,g]	12.54	21.05[g]	11.80	17.20[d,g]
1999	14.88	18.17[g]	14.36	16.24[d,g]	12.94	18.40[g]	12.83	15.61[d,g]
All	13.32%	18.19%[g]	10.47%[d]	11.24%[d,g]	10.11%	22.30%[g]	8.13%[d]	14.10%[d,g]
Trend[e]	.504	.457	.766[f]	.945[f,h]	.655	.039[h]	.782	.707[f]
(p-value)	(.01)	(.01)	(.01)	(.01)	(.01)	(.78)	(.01)	(.01)

(*continued on next page*)

TABLE 2 (continued)

All statistics are reported using quarterly data. Forecast error is computed as earnings per share for the quarter minus the mean forecast of earnings per share as of final month of the quarter. All earnings per share amounts are those reported by I/B/E/S.

[a] *AdjFE* = forecast error from the split-adjusted I/B/E/S historical summary tapes.

[b] *ActFE* = forecast error from the actual (non-split-adjusted) I/B/E/S historical summary tapes.

[c] *Ratio* = (AdjFE/ActFE) − 1.

[d] The percentage for *ActFE* is not equal to the percentage for *AdjFE* at $p < .05$ using a Chi-square test.

[e] The amount reported is the regression coefficient using the percentage of zero forecast errors each year (× 100) as the dependent variable and the year as the independent variable.

[f] The coefficient for *ActFE* is not equal to the coefficient for *AdjFE* at $p < .05$ using an F-test.

[g] The percentage for large (high price-to-book) firms is not equal to the percentage for small (low price-to-book) firms for *AdjFE* or for *ActFE* at $p < .05$ using a Chi-square test.

[h] The coefficient for large (high price-to-book) firms is not equal to the coefficient for small (low price-to-book) firms for *AdjFE* or for *ActFE* at $p < .05$ using an F-test.

consistent with our expectation that as the stock split factor declines as time passes, the differential classification of *AdjFE* and *ActFE* will decline.

Researchers have also been interested in the trend of zero forecast error over time (e.g., Brown 1997, 2001; Matsumoto 2002). We calculate the trend as the estimated coefficient in a regression that uses the percentage of zero forecast errors each year (multiplied by 100) as the dependent variable and the year as the independent variable. The trend coefficient presented in Panel A shows that while there is an increasing trend in the percentage of zero observations for *AdjFE* (.505), the trend is statistically greater for *ActFE* (.821). In addition, while the percentage of zero *AdjFE* levels off around 1992, *ActFE* shows a continuing increase in the proportion of zero observations from 1992 to 1999. Untabulated results for the 1992–1999 period show the trend coefficient is not significant for the adjusted data but it is significant for the actual data. Using the adjusted I/B/E/S data, one may conclude that the increasing pattern of zero forecast errors has leveled off in recent years when, in fact, it has continued to increase.

Panel A of Table 2 also reports the interquartile ranges for analysts' forecast errors by year. The interquartile ranges indicate that not only are the number of observations at zero affected by stock split adjustment, so are the overall distributions. The distribution of forecast errors is noticeably tighter for the adjusted data. The differences in the distributions are particularly noticeable in earlier years.

In Panel B of Table 2 we demonstrate how the adjustment made by I/B/E/S could impact conclusions about the relation between certain firm-specific characteristics and earnings thresholds. Prior research has shown that larger firms and higher price-to-book firms are more likely to meet analysts' forecasts (e.g., Elgers and Murray 1992; Skinner and Sloan 2002). Using data from the annual Compustat files, we sort the sample observations into groups based on their beginning-of-year market value of equity or price-to-book ratio. Those in the bottom third of the distribution are classified as small firms or low price-to-book firms. Those in the top third are classified as large firms or high price-to-book firms. We compare the distributions of adjusted versus actual data amounts for these groups.[11]

The percentage of zero *AdjFE* and *ActFE* for small and large firms is shown in the first four columns in Panel B of Table 2. For both small and large firms, the percentage of observations with zero *AdjFE* is significantly higher than what actually occurred. Comparing *AdjFE* for small and large firms indicates that large firms were more likely to exactly meet analysts' expectation. This pattern holds for all years. However, *ActFE* indicates that larger firms began to consistently report zero forecast error more often than small firms only in recent years (1996–1999). For all years combined, larger firms are more likely to hit analysts' forecasts, but the actual differential (10.47 percent vs. 11.24 percent) is much smaller than the results based on *AdjFE* (13.32 percent versus 18.19 percent). The trend coefficients for *AdjFE* indicate no difference for large and small firms (.457 versus .504), while the coefficients for *ActFE* indicate the trend for large firms is significantly greater than the trend for small firms (.945 versus .766; $p < .05$).

---

[11] The average split-adjustment factor for small (large) firms is 1.53 (2.31). The average split-adjustment factor for low (high) price-to-book firms is 1.51 (2.35). The correlation between size ranking and price-to-book ranking is 18 percent, suggesting that these variables primarily provide a different ranking of the data.

The distributions for low and high price-to-book firms are shown in the next four columns of Panel B, Table 2. The *AdjFE* columns show strong evidence that high price-to-book firms are more likely to report earnings equal to analysts' forecasts. The magnitude of the difference is reduced when observing the *ActFE* data, but it is still significant. Interestingly, the percentage of zero observations for *AdjFE* high price-to-book firms is stable over time and the trend coefficient (.039) is not significantly different from zero. However, when looking at the *ActFE* data, the proportion of high price-to-book firms reporting zero forecast error is increasing over time (.707; $p < .01$). The distribution and trend of *AdjFE* are similar to those of *ActFE* for low price-to-book firms, as these firms tend to have fewer stock splits. While the conclusion that high price-to-book firms are more likely to report zero forecast errors has not changed, the results demonstrate the potential effect that failure to consider the I/B/E/S adjustment procedure can have on conclusions, especially in the earlier years of the data.

Overall, the results reported in Panel B of Table 2 demonstrate differences between adjusted and actual forecast errors that are systematically related to firm characteristics. Larger firms and higher price-to-book firms tend to represent successful, growing firms that are more likely to have stock splits. Thus, these firms are more likely to be affected by rounding split-adjusted data. Researchers should be aware that cross-sectional comparisons of earnings thresholds might be affected by the correlation of numerous firm characteristics and the occurrence of stock splits.

Replication of Prior Research

Distribution of Forecast Errors, Earnings, and Earnings Changes

DeGeorge et al. (1999) investigate managements' incentives regarding reported performance with respect to three earnings thresholds: analysts' forecasts, prior year's reported earnings, and reporting a profit in the current period. They suggest that exactly meeting a threshold (i.e., analysts' forecast error equal to zero) is likely the point at which managers modify reported earnings in a way that allows researchers to detect those actions. DeGeorge et al. (1999) do not attempt to identify the process that management uses to modify reported earnings. Instead, they look for abnormal amounts of reported performance at or around these thresholds as indicators that earnings targets are considered by management. Their results suggest that each threshold is an important determinant of reported earnings.

DeGeorge et al. (1999) test the proportion of observations at each of the thresholds using adjusted I/B/E/S data. Because of the rounding procedure used by I/B/E/S, some of the observations around these thresholds are likely misclassified. This is especially true for observations at each earnings threshold (i.e., those observations with a zero amount). Additionally, DeGeorge et al. (1999) do not "unadjust" the I/B/E/S data by multiplying by the split factor or scaling by some per share amount (e.g., price, earnings per share), which increases the buildup of observations at and around zero (see footnote 4).

To investigate the potential impact of the I/B/E/S rounding procedure on the results reported by DeGeorge et al. (1999), we construct a dataset for the same time period 1985–1996, include only companies with years ending in the months of March, June, September, and December, and remove all observations in the upper and lower decile of stock price to be consistent with their research design.[12] Based on our results in Table 2, we expect that the distributions reported by DeGeorge et al. (1999) are likely affected by the rounding adjustment. Specifically, we expect their analyses to indicate more observations at zero for each threshold and a tighter distribution of observations around zero than actually occurred.

The first two columns of results in Table 3 report the number of observations for \$.01 forecast error intervals, ranging from \$.15 to −\$.15. The distribution for the adjusted forecast error is shown in the first column. As expected, the interval with the greatest number of observations is at zero forecast error. The distribution for the actual forecast errors in the second column is noticeably different. The adjusted data have 11,274 observations with zero forecast errors compared to 7,012 observations for the actual data; this difference is statistically significant at $p < .05$ using a Chi-square test.[13] Compared to the distribution for adjusted forecast errors, fewer firms actually hit their forecasted target. The distribution is also much less peaked around zero forecast error. DeGeorge et al. (1999, Appendix, 30–32) derive a test statistic to determine if the proportion of observations at zero forecast error is statistically significantly different from the proportion of observations at −\$.01. They report a test statistic of 6.61 (DeGeorge et al. 1999, 20) and suggest that a test statistic greater than 2 indicates significance. Using the adjusted (actual) data we calculate a test statistic of 7.40 (4.50).

Of the 11,274 cases where the firm appears to exactly meet its forecast based on the adjusted data, 4,262 (38 percent) are misclassified. The third column of Table 3 shows the forecast errors based on the actual data for the 4,262 misclassifications. The actual forecast error was less (greater) than zero 43.8 percent (56.2 percent) of the time. The actual forecast error ranged from a −\$.09 to +\$.12. Recall, based on the adjusted data, that all of these observations had zero forecast error.

The next two columns in Table 3 report the distributions for adjusted and actual changes in earnings per share. The distribution for actual changes in earnings is flatter at zero and across most intervals. Using the adjusted data, the number of observations that are reported as having no change in earnings is greater than the number of observations that actually had no change in earnings (2,768 versus 1,790, $p < .01$). The DeGeorge et al. (1999) test statistic produces the following results for change in earnings per share: DeGeorge et al. (1999) = 5.63, our results using the adjusted data = 6.53, and the actual data = 5.48. The adjusted data overstate the number of firms that just meet or slightly beat prior year's earnings. Overall, 978 of the 2,768 observations that have zero earnings change in the adjusted data are misclassified. Of the 978 misclassifications, 67.3 percent actually reported positive earnings changes and 32.7 percent had negative earnings changes.

Examination of the final three columns in Table 3 shows the distributions for adjusted and actual earnings per share as reported by I/B/E/S and the distribution of actual earnings for those observations with an adjusted earnings level of zero. The two distributions are relatively similar for the intervals below zero. For the intervals at and above zero, the two

---

[12] To ensure that changes in the reported stock split factor between our dataset and that used by DeGeorge et al. (1999) are not driving our results, we replicated all analyses in this section using I/B/E/S data created in March 1998 and find qualitatively similar results.

[13] Since the frequency of observations in one cell affects the frequency of observations in another cell, the Chi-square tests are not independent across forecast error intervals.

TABLE 3
Number of Quarterly Observations in Each Forecast Error, Earnings Change, and Earnings Interval Using the Procedures of DeGeorge et al. (1999)

	Analysts' Forecast Error			Change in Earnings Per Share			Earnings Per Share		
Interval	**Adj[a]**	**Actual[b]**	**Misclass (Adj = 0)[c]**	**Adj[a]**	**Actual[b]**	**Misclass (Adj = 0)[c]**	**Adj[a]**	**Actual[b]**	**Misclass (Adj = 0)[c]**
< −.15	8972	11121[d]		8170	9772	1	6331	6570[d]	
−.15	560	687[d]		458	509[d]		200	209	
−.14	609	735[d]		449	486		235	231	
−.13	715	834[d]		509	571[d]		200	190	
−.12	773	874[d]		589	602		238	239	
−.11	849	976[d]		573	583		249	242	
−.10	1006	1150[d]		629	671		227	239	
−.09	1198	1288[d]	4	669	708		297	277	
−.08	1268	1367[d]	2	827	756[d]	1	277	259	
−.07	1508	1635[d]	5	901	829[d]	1	289	256	
−.06	1764	1907[d]	13	988	883[d]		290	279	
−.05	2167	2293[d]	12	1116	1003[d]	1	315	300	
−.04	2687	2625	26	1257	1053[d]	6	296	261	
−.03	3302	3166[d]	97	1330	1179[d]	24	341	310	
−.02	4533	4052[d]	347	1560	1184[d]	71	290	263	1
−.01	6273	5339[d]	1359	1653	1210[d]	215	263	235	21
.00	11274	7012[d]		2768	1790[d]		518	431[d]	
.01	8542	6735[d]	1689	3790	2063[d]	431	856	770[d]	60
.02	5484	4923[d]	456	4423	2297[d]	112	841	773[d]	4
.03	3763	3758	149	4099	2431[d]	51	931	707[d]	1
.04	2515	2752[d]	57	3700	2518[d]	23	1007	684[d]	
.05	1974	2295[d]	22	3035	2492[d]	13	1335	876[d]	
.06	1407	1727[d]	13	2492	2438	8	1284	789[d]	
.07	1151	1413[d]	4	1935	2201[d]	10	1533	865[d]	
.08	944	1204[d]	3	1664	2028[d]	3	1647	892[d]	
.09	742	1030[d]	1	1299	1705[d]		1617	804[d]	
.10	676	910[d]	1	1146	1614[d]	2	1907	1174[d]	
.11	526	711[d]	1	1007	1356[d]		1804	974[d]	
.12	460	626[d]	1	854	1282[d]		1874	1108[d]	
.13	376	547[d]		704	1093[d]	1	1942	1084[d]	
.14	354	526[d]		630	945[d]		1958	1074[d]	
.15	313	432[d]		544	857[d]		1843	1271[d]	
> .15	3340	5375[d]		7322	11981	4	49427	58029[d]	
Total	82025	82025	4262	63090	63090	978	82665	82665	87
% < 0	46.6%	48.8%	43.8%	34.4%	34.9%	32.7%	12.5%	12.5%	25.3%
% = 0	13.7%	8.5%		4.4%	2.8%		0.6%	0.5%	
% > 0	39.7%	42.7%	56.2%	61.3%	62.3%	67.3%	86.9%	87.0%	74.7%

[a(b)] Adjusted (actual) amount as reported by I/B/E/S.

[c] Observations classified at zero for the adjusted data, but not classified at zero for the actual data.

[d] Proportion of observations in this interval are significantly different between the adjusted and actual data at $p < .05$ using a Chi-square test. Since the frequency of observations in one cell affects the frequency of observations in another cell, the Chi-square tests are not independent across forecast error intervals.

distributions differ substantially. The adjusted data overstate the number of firms that just meet or slightly beat prior year's earnings (518 versus 431, $p < .01$). The DeGeorge et al. (1999) test statistic produces the following results for earnings per share: DeGeorge et al. (1999) = 3.84, our results using the adjusted data = 3.06, and the actual data = 2.41.

Overall, the analysis reported by DeGeorge et al. (1999) is affected by the rounding procedure, especially for the zero value of each threshold. However, it is important to note that some of the differences are small, particularly for the earnings-level threshold. Also applying the test statistic derived by DeGeorge et al. (1999) to our actual data produces conclusions that are consistent with theirs. However, significant differences clearly exist between the classifications derived from the adjusted versus the actual data. For forecast errors, the adjusted data produces a significantly different classification of observations for 30 of 31 one-cent intervals in Table 3. For earnings changes and earnings levels, the classification is significantly different from −\$.09 to +\$.15 and \$.00 to \$.15, respectively. We recommend using actual data to avoid drawing potentially erroneous conclusions when investigating earnings thresholds, especially in settings where firm characteristics such as size are important (see Panel B of Table 2).

Inferior Returns to Growth Stocks

Skinner and Sloan (2002) investigate the reason for the lower future returns of growth stocks relative to value stocks. They find that the market's negative reaction to negative earnings surprises is more pronounced for growth stocks. However, the market's positive reaction to positive earnings surprises does not differ between growth and value stocks. Skinner and Sloan (2002) argue that growth stocks receive a greater reaction to negative forecast errors because investors were too optimistic in the current period and this over-optimism is corrected in subsequent periods as earnings forecasts are not met. Their results are consistent with the inferior subsequent returns to growth stocks being attributable to expectational errors about future earnings performance, rather than being attributable to risk differences or methodological problems associated with calculating long-term abnormal returns. After controlling for the asymmetric price reaction to negative earnings surprises of growth firms, Skinner and Sloan (2002) find no difference in the subsequent returns of growth and value stocks.

To test their primary hypothesis, Skinner and Sloan (2002) estimate the following model:

$$\begin{aligned} R_{it\tau} = {} & \alpha + \beta_1\, Growth_{it} + \beta_2\, Good_{it\tau} + \beta_3\, Bad_{it\tau} \\ & + \beta_4\, (Good_{it\tau} \times Growth_{it}) + \beta_5\, (Bad_{it\tau} \times Growth_{it}) + \varepsilon_{it\tau} \end{aligned} \tag{1}$$

where:

i = firm index;
t = calendar quarter in which growth portfolio assignments are made;
τ = 20 subsequent quarters over which returns and earnings surprises are tracked;
$R_{it\tau}$ = size-adjusted return for firm i in quarter $t+\tau$, beginning two days after the announcement of earnings in quarter $t+\tau-1$ and ending the day after the announcement of earnings in quarter $t+\tau$;
$Growth_{it}$ = growth quintile to which firm i is assigned in quarter t (0 = low growth quintile,...,4 = high growth quintile);

$Good_{it\tau}$ = indicator variable taking the value of 1 if the firm-quarter observation reports a positive earnings surprise in quarter $t+\tau$, and 0 otherwise; and

$Bad_{it\tau}$ = indicator variable taking the value of 1 if the firm-quarter observation reports a negative earnings surprise in quarter $t+\tau$, and 0 otherwise.

Skinner and Sloan (2002) measure the earnings surprise as I/B/E/S earnings minus the median forecast of earnings in the final month of the fiscal quarter. The coefficient on Growth (β_1) estimates the relation between growth in quarter t and subsequent abnormal returns in quarters $t+\tau$. The coefficients on the interaction terms (β_4 and β_5) estimate the differential price reaction to positive and negative earnings surprises for growth and value stocks, respectively.

Because of the rounding procedure used for the I/B/E/S data, the *Good* and *Bad* variables will be measured with error. Many of the observations that report zero forecast error actually had positive or negative forecast errors. A greater misclassification of positive forecast errors (a likely outcome given the analysis in Table 3) will lead to the average return at zero forecast error being overstated (i.e., too positive) when using the adjusted data. The overstated return at zero forecast error will cause the difference between firms classified as having zero forecast error and those that classified as missing (beating) the forecast to be overstated (understated). Furthermore, the potential for misclassification is greater in the growth portfolio because these firms are more likely to have stock splits (e.g., see Panel B of Table 2). Thus, the adjusted data have the potential to create an appearance of an asymmetric price response for just missing versus just beating the forecast, and the magnitude of the asymmetry could vary with growth. The greater the frequency of stock splits, the disproportionately greater the misclassification of positive forecast errors to zero and the more overstated the return to zero forecast error.

To determine the impact that the rounding procedure has on the results, we estimate Equation (1) using both the adjusted and actual I/B/E/S data. We employ similar sampling procedures, variable measurements, and statistical tests as those in Skinner and Sloan (2002). The results are reported in Table 4. For convenience, we also tabulate the results reported by Skinner and Sloan (2002) (Panel A) along with ours (Panel B).

We first estimate the relation between growth quintiles and stock returns over the subsequent 20 quarters, without controlling for the asymmetric response to positive and negative earnings surprises. Similar to Skinner and Sloan (2002), we find a significantly negative relation between growth and subsequent returns. Next, we control for subsequent earnings news (*Good* and *Bad*) and the asymmetric price response for growth firms (*Good* × *Growth* and *Bad* × *Growth*), where *Good* versus *Bad* are based on the adjusted I/B/E/S data. As in Skinner and Sloan (2002), the coefficient on *Growth* is no longer significant, while the coefficient on the interaction of *Growth* and *Bad* is significant. The negative price reaction to a negative earnings surprise appears more severe for growth stocks. After controlling for this, there is no evidence of inferior returns to growth stocks. The coefficient on the interaction of *Growth* and *Good* is not significant. These results lead Skinner and Sloan (2002) to conclude that the inferior subsequent returns of growth stocks are the result of expectational errors about future performance.

In the final row of Table 4, we estimate Equation (1) using the actual I/B/E/S data. There are two notable differences. First, the coefficient on growth remains significantly negative after controlling for any asymmetric price response. Second, the coefficient on the interaction of *Growth* and *Bad* becomes less significant, going from $t = -5.08$ ($p = .0001$) for the adjusted data to $t = -1.74$ ($p = .0818$) for the actual data. The decrease in significance indicates that the differential market penalty associated with missing a forecast is

TABLE 4

The Relation between Growth and Future Abnormal Returns over the Subsequent 20 Quarters, Controlling for the Asymmetric Response to Good and Bad News

Panel A: Skinner and Sloan's (2002) Results

	Intercept	*Growth*	*Good*	*Bad*	*Good × Growth*	*Bad × Growth*	*Adj. R²*
(Table 2)	*.0065*	*−.0032*					*.06%*
	(5.36)	*(−6.56)*					
(Table 3)	*.0158*	*.0020*	*.0347*	*−.0481*	*.0022*	*−.0095*	*7.57%*
	(3.81)	*(1.48)*	*(7.70)*	*(−10.78)*	*(1.33)*	*(−5.95)*	
Panel B: Our Results							
	.0102	−.0029					.05%
	(7.27)	(−5.03)					
Adjusted	.0250	−.0006	.0326	−.0603	.0018	−.0084	8.44%
I/B/E/S data	(6.27)	(−.43)	(7.26)	(−13.55)	(1.10)	(−5.08)	
Actual	.0235	−.0037	.0334	−.0580	.0051	−.0035	8.21%
I/B/E/S data	(4.63)	(−2.02)	(6.13)	(−10.69)	(2.54)	(−1.74)	

The dependent variable is the size-adjusted return, beginning two days after the announcement of earnings in quarter $t + \tau - 1$ and ending the day after the announcement of earnings in quarter $t + \tau$.

As in Skinner and Sloan (2002), all t-statistics are reported after dividing the initial t-statistic in the pooled model by the square root of 20.

$Growth_{it}$ = growth quintile to which firm i is assigned in quarter t (0 = low growth quintile,...,4 = high growth quintile);

$Good_{it\tau}$ = indicator variable taking the value of 1 if the firm-quarter observation reports a positive earnings surprise in quarter $t + \tau$ and 0 otherwise; and

$Bad_{it\tau}$ = indicator variable taking the value of 1 if the firm-quarter observation reports a negative earnings surprise in quarter $t + \tau$ and 0 otherwise.

not as severe for high- and low-growth stocks as suggested by the adjusted data. The significance of the interaction term *Growth* and *Good* increases from t = 1.10 (p = .2713) for the adjusted data to t = 2.54 (p = .0111) for the actual data. The reaction to good news is greater for high-growth firms. Using the actual I/B/E/S data, we conclude that the inferior returns to growth stocks are not related to expectational errors about future performance.[14]

Relation between Analyst Uncertainty and Future Returns

Prior research has shown a negative relation between the standard deviation of earnings forecasts and future returns (e.g., Ackert and Athanassakos 1997; Ang and Ciccone 2001). Ackert and Athanassakos (1997) argue that the standard deviation of earnings forecasts is a proxy for *ex ante* uncertainty, and analysts tend to be too optimistic when uncertainty is higher. If the market follows analysts' optimism, then prices will also be too high. Ackert and Athanassakos (1997) therefore hypothesize and find that uncertainty (i.e., standard deviation of earnings forecasts) is negatively related to future returns. They argue that this finding would enable them to earn abnormal returns because "this information about forecast dispersion is available to investors at the beginning of the holding period" (Ackert and Athanassakos 1997, 270).

---

[14] Skinner and Sloan (2002, Table 5) also estimate Equation (1) over different return intervals. Our results using the adjusted data for these alternative return windows produce identical conclusions to theirs. Using the actual data, however, we find no evidence across these return intervals of an asymmetric response to bad news and no evidence that the inferior returns to growth stocks are related to expectational errors by the market.

However, the standard deviation of earnings forecast is not totally an *ex ante* measure as reported on the I/B/E/S database; it is also an *ex post* measure because of future stock splits. I/B/E/S reports the standard deviation of earnings forecasts based on adjusted amounts. Therefore, a future stock split will reduce the level and standard deviation of forecasted earnings per share in the current period (i.e., a future 2-for-1 split will reduce the current standard deviation by one-half, a future 3-for-1 split will reduce the current standard deviation by one-third, and so on).[15] Firms with future stock splits are more likely the ones that have higher future returns (Lakonishok and Lev 1987). In addition, firms with future stock splits are more likely to report lower standard deviation of earnings forecasts as reported by I/B/E/S. Putting these together, the prediction that firms with lower (higher) standard deviation of earnings forecasts will have higher (lower) future returns could be mechanically driven because the standard deviation is not purely an *ex ante* measure.

Following the research design in Ackert and Athanassakos (1997), we use data from 1980 to 1991 to classify firms into quartiles each year based on their standard deviation of year *t* earnings forecasts made in June of year $t-1$, and then we measure raw returns and beta excess returns over the subsequent 20-month period beginning in July of year $t-1$ and ending February of year $t+1$. The results are reported in Table 5. The results in Panel B for the adjusted I/B/E/S data are similar to those in Panel A by Ackert and Athanassakos (1997). Firms with higher standard deviation of earnings forecasts have lower future returns. The difference between the low and high quartile's average monthly raw (beta excess) return is .84 percent (.78 percent), or an annualized return of 10.56 percent (9.77 percent). These differences are statistically and economically significant.[16] These findings lead Ackert and Athanassakos (1997, 272) to conclude that they "can construct profitable portfolio strategies" by creating a hedge portfolio that is long (short) in firms that have low (high) standard deviation of earnings forecasts as reported by I/B/E/S in June of each year.

Next we repeat this analysis using the actual I/B/E/S data. In this analysis, the standard deviation is a truly *ex ante* measure. The results reported in Panel C of Table 5 do not support the negative relation between the standard deviation of earnings forecasts and future returns. In fact, the relation between actual standard deviation and future return is slightly positive. The difference between the low and high quartiles is −.05 percent for raw returns and −.14 percent for beta excess returns. While the beta excess return is statistically significant, it is only 1.69 percent on an annualized basis.[17]

The third rows in Panels B and C of Table 5 highlight the source of the differences by showing the average stock split factor for each quartile. In Panel B, there is a significant negative relation between the split factor and standard deviation of earnings forecasts. The

---

[15] It is even possible that a firm with a high enough split factor will report a zero standard deviation, when in fact the standard deviation was positive. For example, suppose a firm had three analysts who made forecasts of $.08, $.09, and $.10. In a subsequent period, if the firm had a 3-for-1 stock split, I/B/E/S would report the three forecasts as $.03, $.03, and $.03 and the standard deviation as $.00. Limiting the sample observations in Tables 1 and 2 to those with three or more analysts following, we find that the adjusted I/B/E/S data shows that 14.1 percent of the observations have zero dispersion (i.e., standard deviation of forecasted earnings). However, only 5.1 percent of the observations actually have zero dispersion.

[16] If we include the data from 1992–1999 in our analysis, the difference in raw returns for the high and low quartiles is .89 percent ($p < .01$) and the difference in beta excess returns is .90 percent ($p < .01$).

[17] If we include the data from 1992–1999 in our analysis, the difference in raw returns for the high and low quartiles is −.24 percent ($p < .01$) and the difference in beta excess returns is −.24 percent ($p < .01$).

TABLE 5
The Relation between Analyst Uncertainty and Future Average Monthly Returns

	Standard Deviation of Earnings Forecast				
	Q1 (Low)	*Q2*	*Q3*	*Q4 (high)*	*Q1–Q4*[a]
Panel A: Ackert and Athanassakos' (1997) Results for the Period 1980–1991					
Raw return	*.0186*	*.0155*	*.0129*	*.0097*	*.0089* *(.01)*
Beta excess return	*.0038*	*.0008*	*−.0017*	*−.0043*	*.0081* *(.01)*
Panel B: Our Results for the Adjusted I/B/E/S Data for the 1980–1991 Period					
Raw return	.0172	.0137	.0110	.0088	.0084 (.01)
Beta excess return	−.0014	−.0049	−.0071	−.0092	.0078 (.01)
Split factor	8.105	2.825	2.001	1.378	6.727 (.01)
Panel C: Our Results for the Actual I/B/E/S Data for the 1980–1991 Period					
Raw return	.0124	.0132	.0121	.0129	−.0005 (.41)
Beta excess return	−.0065	−.0051	−.0057	−.0051	−.0014 (.02)
Split factor	3.419	3.727	3.399	3.719	.300 (.16)

[a] The amounts shown are the differences in means between the lowest and highest quartile. p-values of the two-tailed t-test that the difference equals 0 are shown in parentheses.

Following Ackert and Athanassakos (1997), in each year firms are sorted evenly into quartiles based on the standard deviation of year t earnings forecasts in June of year $t-1$. Only firms with December year-ends and with at least three analyst forecasts are included. The return is average monthly return over the 20-month period beginning in July of year $t-1$ and ending in February of year $t+1$. Beta excess returns are computed by compounding for the month the daily return less the equally weighted return of a similar beta portfolio.

lowest (highest) quartile of standard deviation has an average split factor of 8.105 (1.378). This is precisely what we would expect to find for adjustment factors greater than 1. Firms that have more stock splits tend to have lower adjusted earnings amounts and therefore lower standard deviations of forecasts. In addition, firms that are more likely to have splits are more likely the ones that have higher stock price performance. As shown in Panel C, the relation between the *ex ante* measure of actual standard deviation and the stock split factors is no longer statistically significant. The lowest (highest) quartile has an average split factor of 3.419 (3.719). Researchers must be aware that the standard deviation reported in the usual I/B/E/S database is affected by future stock splits and therefore is partially an *ex post* measure of stock performance.

Prior research has also documented a positive relation between uncertainty and analysts' forecast optimism (e.g., Ackert and Athanassakos 1997; Das et al. 1998). We find that this relation is reduced (but still statistically significant) when using the actual data instead of the adjusted data. The rank correlation between the standard deviation of earnings forecasts and forecast error is $-.2876$ ($p < .01$) for the adjusted data and reduces to $-.1245$ (p

$< .01$) for the actual data. But, the results in Table 5 suggest that the market understands that analyst uncertainty is positively related with analyst optimism and it adjusts analysts' forecasts appropriately in setting expectations to prevent a relation with future returns.

Market Reactions to On-Target Earnings Announcements

Baber and Kang (2002) point out that split adjusting and rounding analyst forecast data can result in "contaminated" observations (i.e., observations reported as having had zero forecast error but actually had nonzero forecast error). Since splits occur for superior performers *ex post*, Baber and Kang (2002) speculate that the contaminated observations are predominantly positive forecast errors that likely have positive announcement period returns. Inclusion of these contaminated observations therefore overstates the announcement period returns of on-target firms. They do not use actual forecast data to test their speculation but instead rely on the association between announcement period returns and split factors. Baber and Kang (2002) find that on-target firms with larger split factors tend to have more positive announcement period returns. They interpret this as evidence of spurious results caused by rounding adjusted forecast data.[18]

Using sampling procedures and variable measurements similar to Baber and Kang (2002), we replicate their study by observing the three-day abnormal return surrounding on-target earnings announcements for firms sorted by split factor.[19] The first two columns in Table 6 present Baber and Kang's (2002) results. As shown in the next two columns, our replication of their study using the adjusted I/B/E/S data produces a similar positive relation between announcement period returns and split factors. Mean abnormal returns are negative (positive) in the three lowest (highest) split factor categories. In the final four columns of Table 6, one can see that approximately 70 percent of contaminated observations [= 1,717/(721+1,717)] had positive forecast errors and a positive average announcement period return of 1.17 percent. These results are consistent with Baber and Kang's (2002) speculation that contaminated observations introduce a spurious positive stock performance into the analysis of on-target observations.

Next, we examine this relation using the actual data. Because we use actual forecast data, we can more directly assess the extent to which the positive relation between announcement period returns and split factors is due to contamination. Using the actual data, we find that the positive relation between split factors and announcement period returns remains. The average return is −.75 (−.39 percent) for observations with split factors less than (equal to) 1 versus .63 percent for observations with split factors greater than 4.[20] These results suggest that the complexity of the relation between on-target earnings announcements and returns extends beyond contamination. While it is beyond the scope of this paper, a number of firm and investor characteristics could explain why some on-target earnings announcement returns are positive while others are negative. The split factor (which is an *ex post* measure) proxies for one or more of these characteristics. For example, it could be that when firms announce earnings for the current period, they also release information about future earnings. Firms that have zero forecast error and reveal positive

---

[18] Baber and Kang (2002) use data provided by First Call. We replicate their reported results using the I/B/E/S adjusted data and derive their conclusions. Their data source has a similar rounding issue (see footnote 2).

[19] As in Baber and Kang (2002), abnormal returns are measured using the market model and the equally weighted index. The market model parameters are estimated over a 150-day period before the earnings announcement (day −170 to day −21).

[20] The rank correlation between the mean split factor and mean announcement period returns for the six portfolios in Table 6 is .9429 ($p = .0048$). The correlation between the split factor and announcement period return for individual observations is .0503 ($p = .0001$).

TABLE 6
Earnings Announcement Returns for Firms that Report Zero Analyst Forecast Error

	Baber and Kang's (2002) Results (Figure 1)[a]		**Our Results**							
							Adjusted FE = 0			
	FE = 0		**Adjusted FE = 0**		**Actual FE = 0**		**Actual FE < 0**		**Actual FE > 0**	
	n	*ARET*	**n**	**ARET**	**n**	**ARET**	**n**	**ARET**	**n**	**ARET**
Factor < 1	*22*	*−2.32%*	67	−.75%	67	−.75%				
Factor = 1	*8,463*	*−.30%*	7,644	−.39%	7,644	−.39%				
1 < Factor ≤ 2	*2,122*	*−.02%*	2,751	−.01%	1,849	−.12%	332	−.29%	570	.54%
2 < Factor ≤ 3	*375*	*.58%*	784	.65%	429	.32%	114	.38%	241	1.37%
3 < Factor ≤ 4	*393*	*.85%*	864	.32%	313	.17%	137	−.99%	414	.87%
4 < Factor	*216*	*1.46%*	811	1.20%	181	.63%	138	−1.11%	492	2.05%
Factor >1	*3,106*	*.27%*	5,210	.34%	2,772	.03%	721	−.47%	1,717	1.17%
Full Sample	*11,591*	*−.15%*	12,921	−.10%	10,483	−.28%				

---

[a] Baber and Kang (2002) use forecast data provided by First Call. These data are provided in a similar format to the I/B/E/S Summary data.

Adjusted FE = 0 represents the sample of firms that report zero forecast error on the typical I/B/E/S database. Forecast error is computed as earnings per share for the quarter minus the mean forecast of earnings per share as of final month of the quarter;

Actual FE = 0 represents the sample of firms that actually reported zero forecast error. The difference between the sample for *Adjusted FE* = 0 and *Actual FE* = 0 represents firms that have been misclassified (i.e., firms that did not report zero forecast error but were reported by I/B/E/S as having zero forecast error);

Factor = the split adjustment factor reported by I/B/E/S for the final month of the quarter; and

ARET = abnormal returns = the three-day period surrounding the quarter's earnings announcement date (−1, 0, and 1), computed as market model excess returns using the equally weighted market index. The parameters of the market model are estimated over the 150-day period preceding the earnings announcement (day −170 to day −21). This is the abnormal return measure used in Baber and Kang (2002).

future earnings news will likely receive a more positive market reaction than will firms that have zero forecast error but release no (or negative) future earnings news. Firms that release positive future earnings news are also the firms that will, on average, perform better in the future and have future stock splits. Therefore, there is some reason to expect a positive relation between stock split factors and current announcement period returns. Eliminating all firms that have factors greater than 1 to "decontaminate" the sample may bias results and lead to incorrect inferences because of the elimination of better performing firms.

In addition to testing for a stock price reaction at the time earnings are announced, researchers may be interested in testing for subsequent price movements. For example, the negative average announcement period returns to on-target observations in Table 6 may lead to speculation that the market has not adequately rewarded the company for meeting its forecasts. In this case, one would predict positive future abnormal returns for on-target firms as the market eventually provides these rewards. However, firms that have higher stock price performance are more likely to have stock splits, and firms with more stock splits are more likely to have prior forecast errors misclassified at zero. We test whether the I/B/E/S rounding procedure could lead to a mechanical relation between higher future returns and current zero forecast errors. This is similar to our analysis in the prior section concerning the association between future returns and forecast dispersion, except that forecast error is used in place of forecast dispersion.

Using the sample and variables generated from our replication of Skinner and Sloan (2002), we measure the future returns of firms that report zero forecast error. Future returns are measured over the four quarters following the announcement of a zero forecast error. For the adjusted I/B/E/S data, we find an average quarterly abnormal return of 1.59 percent (6.51 percent on an annualized basis). In contrast, when we use the actual I/B/E/S data to identify firms that actually had zero forecast error, we find an average quarterly return of only −.07 percent (−.28 percent on an annualized basis). The average quarterly return for all firms in the sample is .04 percent. With the adjusted I/B/E/S data, firms with zero forecast errors appear to earn positive future abnormal returns. But this potential trading strategy is an illusion because, in reality, firms with zero forecast error had future abnormal returns close to zero.

III. CONCLUSION

The forecast data provided by I/B/E/S to the academic research community play an important role in research investigating issues related to financial accounting information. I/B/E/S traditionally provides these data on a split-adjusted basis, presumably to provide comparable time-series data for its customers. For academic researchers, it is more beneficial in many cases to know the actual historical amounts that are not adjusted for subsequent stock splits. This paper investigates potential ramifications to academic researchers from using I/B/E/S split-adjusted data versus data that have not been split-adjusted.

The problem for academics is that after I/B/E/S divides an earnings amount by the split factor, the result is rounded to two decimal places. When decimal places are removed by rounding, the adjustment cannot be reversed. In those cases, the researcher cannot determine prior years' actual amounts. The rounding procedure is especially problematic when analyzing cases where forecast errors equal zero. The typical (adjusted) I/B/E/S database can report a zero forecast error when, in fact, the error based on the unadjusted data is non-zero. No type of adjustment by the researcher can correct this misclassification.

We obtain actual (unadjusted) data from I/B/E/S. We then replicate prior research using the adjusted and actual I/B/E/S data and find that some conclusions are affected. Research conclusions are more likely affected by the rounding procedure in samples that have stock splits (e.g., larger firms, higher price-to-book firms, better performers, etc.) and where the research question focuses on zero amounts (e.g., assessing the percentage of zero forecast errors over time; relating firm characteristics to the probability of zero forecast error; calculating the market's reaction to zero forecast error; inferring earnings management based on the distribution of earnings, earnings changes, and forecast errors around zero). The rounding procedure also reduces variation in forecasts across analysts, which results in a downward bias in forecast dispersion for all firms with subsequent stock splits. The bias will be an increasing function of the split factor, and again, the researcher using only adjusted data cannot fully eliminate this bias.

To ensure that future research findings are not affected by this potential misclassification, we recommend that researchers request the actual data from I/B/E/S and create their own split-adjusted data without rounding to the nearest penny. If the actual data cannot be obtained, an alternative is to recalculate I/B/E/S consensus statistics using the detail I/B/E/S adjusted data. The detail data are rounded to four decimal places, which allows a more accurate recalculation of the actual amount using the split factor provided by I/B/E/S. Research that uses the adjusted summary data from I/B/E/S should present sensitivity test results that include only observations with a stock split factor equal to 1 to ensure the reported results are not affected by the rounding procedure. Additional analyses that exclude observations with stock split factors greater than 2 or 3, large market value firms, or high price-to-book firms should also be considered.

Our results suggest that some conclusions of published research are affected by using the adjusted I/B/E/S data. Furthermore, other research studies may have been abandoned or rejected because the adjusted I/B/E/S data led to statistically insignificant results when, in fact, a relation existed. Using forecast data that are not rounded will increase a researcher's ability to investigate and draw conclusions on issues of importance to the academic and practice community.[21]

REFERENCES

Abarbanell, J., and R. Lehavy. 2001. Differences in commercial database reported earnings: Implications for inferences concerning analyst forecast rationality, the association between prices and earnings, and firm reporting discretion. Working paper, The University of North Carolina at Chapel Hill.

Ackert, L., and G. Athanassakos. 1997. Prior uncertainty, analyst bias, and subsequent abnormal returns. *The Journal of Financial Research* 20 (Summer): 263–273.

Ang, J., and S. Ciccone. 2001. Analyst forecasts and stock returns. Working paper, Florida State University and University of New Hampshire.

Baber, W., and S. Kang. 2002. The impact of split adjusting and rounding on analysts' forecast error calculations. *Accounting Horizons* 16 (December): 277–289.

Brown, L. D., R. L. Hagerman, P. A. Griffin, and M. E. Zmijewski. 1987. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting & Economics* 9 (April): 61–88.

———. 1993. Earnings forecasting research: Its implications for capital markets research. *International Journal of Forecasting* 9 (November): 295–320.

———. 1997. Analysts' forecast errors: Additional evidence. *Financial Analysts Journal* 53 (November/December): 81–88.

———. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39 (September): 221–241.

DeGeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72 (January): 1–33.

Das, S., C. Levine, and K. Sivaramakrishnan. 1998. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review* 73 (April): 277–294.

Elgers, P., and D. Murray. 1992. The relative and complementary performance of analyst and security-price-based measures of expected earnings. *Journal of Accounting & Economics* 15 (June/September): 303–316.

I/B/E/S. 2000. *Research Bibliography.* New York, NY: I/B/E/S.

Lakonishok, J., and B. Lev. 1987. Stock splits and stock dividends: Why, who, and when. *The Journal of Finance* 42 (September): 913–932.

Matsumoto, D. A. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (July): 483–514.

Skinner, D. J., and R. G. Sloan. 2002. Earnings surprises, growth expectations, and stock returns: Don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7 (June/September): 289–312.

---

[21] Of course, even the actual Summary data provided by I/B/E/S will have some rounding effects. Each individual analyst may round his/her forecast of earnings to the nearest penny after dividing by total shares. Then, aggregation of all analysts' forecasts into a single consensus measure may involve additional rounding. Companies also round when reporting earnings share. These rounding issues should not be as significant as those discussed in the paper. There is no reason to believe these would produce a systematic bias.

THE ACCOUNTING REVIEW
Vol. 78, No. 4
2003
pp. 1069–1095

Control in a Teamwork Environment—The Impact of Social Ties on the Effectiveness of Mutual Monitoring Contracts

Kristy L. Towry
Emory University

ABSTRACT: This study examines control in a teamwork setting, experimentally investigating two financial incentive systems that have been proposed in the agency-theory-based analytic literature. Both systems rely on mutual monitoring—the ability of team members to observe each other's actions. However, the systems differ on whether team members report observations of their peers' efforts to management (vertical incentive system) or directly control the actions of each other (horizontal incentive system). Findings suggest that the effectiveness of these systems depends on the level of team identity. Specifically, a strong team identity leads to greater coordination. The result is that the effectiveness of a vertical incentive system is degraded by a strong team identity. On the other hand, a horizontal incentive system becomes more effective in the presence of a strong team identity. The results of this study suggest that when the team has achieved a high level of identity, the most effective way to use this information is likely horizontal in nature, delegating responsibility for control to self-managed teams, rather than extracting the information through reporting mechanisms. This study thus helps explain why firms have more readily embraced horizontal incentive systems than vertical incentive systems.

Keywords: *control systems; teams; mutual monitoring; incentives; social identity.*

Data Availability: *The experimental data are available for purposes of replication. Please contact the author.*

This paper is based on my dissertation at The University of Texas at Austin. I gratefully acknowledge the many contributions of my committee: Steven Kachelmeier (chair), Urton Anderson, Shane Dikolli, Lisa Koonce, Art Markman, and David Platt. I am also grateful for research assistance from Angela Coletti, Jessen Hobson, and Alex Yen, and for helpful comments and suggestions from Rob Bloomfield, Sarah Bonner, Tom Dyckman, Joe Fisher, Sue Haka, Jim Hesford, Ron King, William Kinney, Molly Mercer, Mark Nelson, Paul Newman, Mark Peecher, Bill Rankin, Karen Sedatole, Lisa Sedor, Geoff Sprinkle, Wim Van der Stede, Igor Vaysman, Mark Young, and workshop participants at Cornell University, Emory University, Harvard University, Indiana University, Stanford University, Texas A&M University, The University of Alabama, University of Connecticut, University of Georgia, University of Illinois, University of Notre Dame, University of Southern California, The University of Texas, University of Washington, and Washington University. I appreciate the financial support of the McCombs School of Business Center for Business Measurement and Assurance Services, the Institute of Internal Auditors Research Foundation, the Deloitte & Touche Foundation, and the Eugene and Dora Bonham Memorial Fund.

Editor's note: This paper was accepted by S. Jane Kennedy, Editor.

Submitted October 2002
Accepted April 2003

I. INTRODUCTION

The use of teams in the workplace has increased dramatically over the last several years (Cohen and Bailey 1997). Firms report a number of benefits of teams, including greater participation and involvement, increased attention to process improvements, and improved employee satisfaction (Wellins et al. 1994). However, from a measurement and contracting perspective, the teamwork setting presents a challenge. When accurate and verifiable measures of individual contributions are not available, incentive contracting may not be useful for encouraging employees to direct their efforts toward management's goals. Instead, the use of team incentives can result in free-riding. Therefore, researchers have begun to design incentive systems tailored to the teamwork setting.

This study investigates two incentive systems. Both systems rely on mutual monitoring, the ability of team members to observe each other's actions. Peer observations are "soft" in the sense that they are not directly verifiable and do not meet the stringent measurement requirements to be included in the accounting library (Demski 1997).[1] Nevertheless, incentive contracting schemes can incorporate peer observations, and analytic researchers have identified two general approaches for doing so. The approaches differ on whether team members report observations of their peers' efforts to management (vertical incentive system) or directly control the actions of each other (horizontal incentive system). The premise of this study is that the effectiveness of each system depends on the degree to which team members have established a strong psychological attachment, or *team identity*. Specifically, I show that team identity *interacts* with the incentive system, such that as the level of identity increases, the horizontal approach becomes more effective relative to the vertical approach.

This interaction occurs because of key assumptions underlying the two types of incentive systems. A key feature of agency models using the vertical approach is the assumption that the agents will choose their strategies independently, rather than engaging in coordinated behavior. Specifically, a vertical incentive system reduces a worker's incentive to shirk, because of the threat that his/her behavior will be observed and reported by a teammate. This threat can be eliminated if all agents agree to "cover for each other," consistently making favorable reports regardless of the effort levels observed. Horizontal systems, on the other hand, rely on the agents' coordination. That is, the principal creates an incentive system that induces the agents to agree (among themselves) to the high effort levels desired by the principal and to enforce these agreements through the use of formal sanctions, peer pressure, or enforceable side-contracting. I demonstrate that team identity increases the level of coordination among team members, thereby decreasing the effectiveness of the vertical system, while increasing the effectiveness of the horizontal system.

This research is important to managers wishing to maximize the effectiveness of work teams and to managerial accountants and assurance providers who design and evaluate managerial control systems. A considerable body of analytic modeling research has focused on the use of incentives to motivate desired behavior by employees. Milgrom and Roberts (1992, 413) point out, however, that:

> Almost all of the formal theory emphasizes incentives for individuals on the grounds that it is individuals who must be motivated to work. Yet the most common explicit incentive contracts are applied across groups of individuals.

---

[1] Compensation contracts frequently allow for the use of soft data, generally through subjective performance evaluations (Fisher et al. 2003) or discretionary bonuses (Baiman and Rajan 1995).

Thus, more recent analytic work has focused on incentive systems specifically designed for the teamwork environment. This research contributes to this stream of literature by investigating the use of an unconventional measure (peer observation) in incentive contracting. Further, this study provides evidence that the incentive system cannot be considered in a vacuum. While the analytic literature tends to view a team as a collection of individuals, united only by interdependent tasks or incentives, this paper shows that social psychological factors—specifically, the strength of team identity—play an important role in determining the effectiveness of incentive systems. In doing so, this paper answers recent calls for research that melds theory from both economics and psychology to provide insights into accounting issues (Haynes and Kachelmeier 1998; Moser 1998; Waller 1995, 2001).

Graduate business students participated in an experiment with a 2 × 2 (Incentive System × Team Identity) between-subjects factorial design.[2] Each experimental session involved eight participants, randomly assigned to four two-person teams. Team members assumed the roles of two division managers working for the same company, with compensation dependent upon their decisions. Each person's primary responsibility was to choose the level of effort (or more broadly, the level of resources) his/her division would commit to the production process. The first manipulated factor is the incentive system, capturing the two models of interest (vertical versus horizontal). The second factor is team identity, manipulated by varying the salience of intra-team relationships.

The primary dependent variable is the level of costly effort chosen. I find a significant interaction between the incentive system and the level of team identity. Specifically, under the horizontal approach, a strong team identity leads to increased effort levels. However, under the vertical approach, a strong team identity leads to decreased effort levels. Further tests provide evidence that this effect is a result of both a cognitive change, leading team members to focus on joint versus individual outcomes, and an increase in communication. Moreover, the incentive systems themselves differentially reinforce team identity in that, after repeated interaction, the level of team identity becomes higher for teams using the horizontal incentive system than for those using the vertical incentive system.

The remainder of this paper is organized as follows. The following section reviews the relevant literature and introduces a model of the team setting that is used to motivate the hypotheses. Section III describes the methodology. Section IV provides data analysis, and the final section summarizes and concludes the paper.

II. THEORY AND HYPOTHESES

The Use of Teams in the Workplace

Modern organizations increasingly use a team-based approach (Cohen and Bailey 1997). For example, based on their 1995 research on salary growth, Hewitt Associates (1995) report that 65 percent of surveyed companies use teams in some capacity. Similarly, in its *1999 Practice Analysis*, the Institute of Management Accountants (IMA 1999) reports that 73 percent of survey respondents work at companies where management accountants are involved in cross-functional teams. Many young companies (e.g., Southwest Airlines [Kelley 2000], Whole Foods [Fishman 1996], Cisco Systems [Gillmor 2000]) use teams as a central organizing principle, but even older, more established companies herald the benefits of teamwork. For example, General Mills, Inc. has reported that productivity is as much as 40 percent higher in factories using a team-based approach as compared to a traditional approach (Dumaine 1990).

---

[2] The experiment was repeated 20 times, and thus period number is a third (within-subjects) factor. However, the primary hypotheses do not involve period-by-period analysis.

Agency Theory in a Teamwork Environment

The analytic literature on contracting has begun to explore the implications of team settings. Most of the models are based on the principal-agent paradigm, with the basic model (Holmstrom 1979) modified to include multiple agents. While the literature has identified several ways in which the principal can improve his/her contracting position in the multiagent setting, this study focuses on mutual monitoring.[3] That is, while the principal cannot observe the agents' actions, the agents may be able to observe each other. The principal can benefit by designing incentive schemes that exploit this capability.

In this section, I introduce two incentives systems that rely on mutual monitoring. For simplicity, I describe these incentive structures using the parameters from the actual experiment. (For a more technical discussion, including the constraints on parameter choices, see the Appendix.) I define a team as two individuals (agents) engaged in some type of joint production. The principal cannot directly observe the agents, and there is no conventional measure of individual performance, but the two agents can mutually monitor each other's actions. Each agent selects an effort level (shirk or work), and joint output is a stochastic function of these effort choices.[4]

I assume that the two agents are risk-neutral and effort-averse. The analytic papers (Arya et al. 1997; Ma 1988) used as a basis for this model assume risk aversion. In fact, if the agents were risk-neutral (with no bankruptcy constraints), there would be a trivial solution to the problem (selling the firm to the agents). In this experimental setting, however, the agents' risk preferences are not related to the hypothesized effects, and therefore, I assume risk neutrality for ease of exposition and operationalization.[5] In the laboratory, I control for risk preferences by basing experimental payments on the expected values of the stochastic distributions. To operationalize the concept of effort aversion, participants who work are charged 10 points, and participants who shirk are not charged any points.

Next, I will explore the use of two types of incentive systems to induce the agents to work in this setting. As will be explained, a key difference between horizontal and vertical incentive systems is the degree to which agents are assumed to coordinate their strategies. In the "Hypothesis" section, I will argue that teams with strong team identities tend to coordinate their strategies, which works in favor of the horizontal incentive system.

---

[3] These improvements include the use of relative performance evaluations (Arya and Glover 1996a; Frederickson 1992; Gibbons and Murphy 1990; Holmstrom 1982; Mookherjee 1984), improved risk-sharing (Ramakrishnan and Thakor 1991; Villadsen 1995), and gains from synergy and cooperation (Itoh 1991; Tirole 1988; Villadsen 1995).

[4] A key feature absent in this experimental setting is synergy. Real-world teams engaged in joint production are likely to experience gains from synergy. However, the use of mutual monitoring is not affected by whether such gains exist. Further, I am not aware of a theory that would suggest an interaction between the degree of synergy and either of the independent variables in this study. Therefore, to simplify the experimental setting, the conditional probability distribution used here is linear and additive in effort, incorporating no synergies into the production process.

[5] From a strictly economic perspective, if the agents were risk-averse, then the principal would prefer the vertical system, as it imposes no risk on the agents. Thus, in this setting with risk-neutral agents, I avoid making any generalizations regarding the relative desirability of the two systems. Instead, my focus is on the impact of social identity on incentive effectiveness. I am not aware of a theory that would suggest an interaction between risk preferences and social identity, and therefore, inducing risk neutrality should not affect the generalizability of the results.

Vertical Incentive System

Agency theorists have developed an extensive literature on incentive contracts that rely on vertical communication (agent to principal) in a multiagent setting (Arya and Glover 1996b; Demski and Sappington 1984; Demski et al. 1988; Fischer and Hughes 1997; Ma 1988; Ma et al. 1988). While the assumptions and forms of these vertical incentive contracts vary, the general model is one in which each agent observes the other agent's action and truthfully reports it to the principal.[6] Each agent's compensation is then based on the report filed by his/her teammate. The agents are essentially whistle-blowers, who pass along to the principal any information they gather about their coworkers. The vertical approach relies on the assumption that the agents will choose strategies independently, because coordination among them can undermine the reliability of their reports.[7]

Peer evaluations are commonly used for career development purposes. For example, many firms have adopted "360 evaluation" systems, in which each worker evaluates the performance of superiors, subordinates, and peers (Edwards and Ewen 1996). However, there has been a reluctance to report these evaluations upwardly or to tie compensation to them (Antonioni and Park 2001; Jackson and Greller 1998; Thatcher 1996; Wells 1999).[8] In fact, several authorities (e.g., W. Edwards Deming, Peter Drucker, and Tom Peters) warn expressly against doing so (Coates 1998). Thus, the use of vertical reports of peer observations for incentive contracting purposes, as conceptualized in economic theory, is not common in practice.

The vertical incentive system in this study is based loosely on the model proposed by Ma (1988) (hereafter Ma). Under this scheme, each agent's wages have two additive components—*effort pay* and a *reporting bonus/penalty*. An agent's effort pay depends entirely on the report filed by his/her teammate, and equals 20 (0) points if the teammate reports that the agent worked (shirked). Thus, assuming that the teammate reports truthfully, the agent is better off working than shirking. Working will net the agent 10 points (20 points of effort pay minus a 10-point deduction for effort aversion). Shirking will net the agent 0 points (0 points of effort pay with no deduction for effort aversion).

This incentive system depends crucially on truthful reporting, and the reporting bonus/penalty provides the incentive for each agent to tell the truth. Specifically, if an agent reports that his/her teammate shirked, and that report is truthful, then the agent receives a reporting bonus of 10 points. However, if that report turns out to be false, then the reporting agent is penalized 20 points. The principal cannot directly observe the actions of the agents. Therefore, one might ask how the principal can verify the veracity of the agents' reports, paying bonuses to accusing agents only if their accusations are proved truthful. One method (similar to that proposed by Ma) is the use of lotteries. Specifically, in accusing another agent of shirking, the accusing agent is required to accept an output-based lottery that is valuable only if the other agent did in fact shirk. Another alternative is for the firm to employ an auditor. The auditor is called only if one agent accuses the other of shirking.

---

[6] A number of the papers in this category describe a setting of correlated private information rather than unobservable action. This setting is conceptually similar to a setting in which each agent imperfectly observes the other's action.

[7] The defining feature of all vertical systems is that the principal uses the agents' truthful reports to extract rents from the agents. If the agents collude in their reports, then they keep these rents, reducing the principal's wealth and rendering the incentive system ineffective.

[8] Recently, several consulting firms have begun to market their peer evaluation systems for compensation purposes, and businesses have explored tying pay to peer reports. It is unclear, however, whether these implementations will be successful (Edwards and Ewen 1996), and this remains a subject of controversy among Human Resources professionals.

The accusing agent is then rewarded (penalized) if the auditor determines that the accusation is truthful (false). In this study, I do not specify a method for ensuring truthful reporting, but instead generalize this aspect of the model by having the experimenter verify the reports.

An important feature of the vertical incentive system is that verification only occurs when one agent accuses the other of shirking. Verification processes are costly, and so the principal would not want to verify all reports.[9] From the principal's perspective, the optimal approach is to verify only negative reports, because under this approach, the costly verification only occurs off-equilibrium. Therefore, the principal uses the agents' ability to mutually monitor one another to achieve a first-best solution.

The vertical incentive scheme is summarized in normal form in Exhibit 1, Panel A. To make economic predictions using this form, one must first recall the sequential nature of the system. That is, the vertical system involves a two-stage game, with agents making effort choices in stage 1 and reporting choices in stage 2. At the beginning of stage 2, each agent has perfect knowledge of the effort choices made in stage 1. In choosing their reporting strategies, the agents therefore play one of four subgames, depending on what effort levels were selected in stage 1. For example, if Agent 1 works and Agent 2 shirks, then the agents will play the stage 2 subgame represented by the four bottom left squares of the matrix. The unique Nash Equilibrium of each of these stage 2 subgames involves truthful reporting. (The equilibria of the four subgames are highlighted.) Through backward induction, the stage 1 game economically reduces to a choice among these four equilibria. This reduced stage 1 game is summarized in Exhibit 1, Panel B. The unique Nash Equilibrium of this game is for both agents to work. As demonstrated by this backward induction process, the anticipation of truthful reporting in stage 2 (the dominant strategy) assures that both agents work in stage 1.

Economic theory predicts the unique subgame perfect Nash Equilibrium that agents facing this incentive system will choose to work and to report truthfully. However, an examination of Exhibit 1 reveals that this equilibrium is not Pareto optimal to the agents. Both agents are better off if they coordinate their actions, each shirking and then falsely reporting that the other has worked. This collusive outcome is not a Nash Equilibrium, because both agents have strict incentives to report truthfully.[10] However, under some circumstances agents may be able to achieve that outcome. This idea will be explored shortly, but first I will introduce the second type of incentive system—the horizontal incentive system.

Horizontal Incentive System

A second way in which the principal can take advantage of the opportunity for mutual monitoring is through the use of a horizontal incentive system, relying on team self-management and peer-based control (e.g., Arya et al. 1997; Barron and Gjerde 1997; Itoh 1993; Kandel and Lazear 1992; Prendergast 1999; Radner 1986; Ramakrishnan and Thakor 1991; Tirole 1988; Varian 1990; Villadsen 1995). In contrast to the vertical approach, the horizontal incentive system does not involve reporting to the principal. Instead, the principal assumes that the agents will explicitly or implicitly coordinate their actions. Therefore, the principal creates an incentive system that induces the agents to agree (among themselves) to the actions desired by the principal and to enforce these agreements through the use of

---

[9] In fact, if the principal were willing to incur the cost of verifying all reports, then peer monitoring would be of no value.

[10] The term "collusion" is defined as coordination among the agents that is not in the principal's best interest.

EXHIBIT 1
The Vertical Incentive System

Panel A: Normal Form Representation of Vertical Incentive System[a]

Agent 1	**Agent 2** *s,s*	*s,w*	*w,s*	*w,w*
*s,s**	10, 10**	30, 0	−20, 0	0, −10
s,w	0, 30	20, 20	0, 20	20, 10
w,s	0, −20	20, 0	−30, −30	−10, −10
w,w	−10, 0	10, 20	−10, −10	10, 10

---

* Represents the effort choice (made in stage 1) and reporting choice (made in stage 2), respectively (s = shirk, w = work). Note that in stage 2, each agent reports on *the other agent's* stage 1 choice.

** Represents the payoff to Agent 1 and Agent 2, respectively.

Panel B: Normal Form Representation of Reduced Stage 1 Game[b]

Agent 1	**Agent 2** *s*	*w*
*s**	10, 10**	0, 20
w	20, 0	10, 10

---

* Represents the effort choice (made in stage 1) and reporting choice (made in stage 2), respectively (s = shirk, w = work).

** Represents the payoff to Agent 1 and Agent 2, respectively.

[a] Note that the vertical incentive system involves a two-stage game, with agents choosing effort levels in stage 1 and reporting in stage 2. The shaded cells represent the unique Nash Equilibria of the four stage 2 subgames. For example, if both agents shirk, then the agents play the stage 2 subgame represented in the top left four cells of the matrix. The unique Nash Equilibrium of this subgame is for each agent to report that the other agent shirked. That is, given that both agents have shirked, the dominant strategy for each agent is to report truthfully. Truthful reporting is the unique Nash Equilibrium for all four stage 2 subgames, as represented in the shading above.

[b] The game implied by the vertical incentive systems is solved through backward induction. The stage 1 game is reduced to a choice among the Nash Equilibria of the four stage 2 subgames. Thus, each cell in Panel B represents one of the shaded cells in Panel A. The unique Nash Equilibrium in this reduced stage 1 game is for each agent to work. That is, given that each agent will report truthfully in stage 2, each agent's dominant strategy is to work in stage 1.

formal sanctions, peer pressure, or enforceable side-contracting (either explicit or implicit).[11]

Unlike the vertical approach, the horizontal approach corresponds to a practice commonly observed in the real world. The popular press has discussed the role of peer pressure and social norms in helping Japanese companies achieve success (Nahavandi and Aranda 1994). Likewise, North American firms have shifted control from supervisors to self-managing teams (Dumaine 1990), often using team-based compensation (DeMatteo et al. 1998) and lateral control regimes (Lazega 2000).

---

[11] While the enforcement mechanism differs across models, the defining characteristic of all horizontal incentive systems is that the principal's preferred outcome requires coordination among the agents. Therefore, as the level of coordination increases, the incentive system becomes more effective.

The horizontal incentive system presented here is based on that of Arya et al. (1997) (hereafter AFG). AFG's approach is to have the principal tie each agent's pay to the team output. A fundamental issue with team-output-based pay is the opportunity for social loafing (the withholding of effort as individual effort becomes less identifiable) (Kidwell and Bennett 1993). To prevent social loafing, AFG rely on the ability of the agents to monitor each other's efforts and to punish each other for shirking. This threat of punishment is a type of peer pressure, as conceptualized by Kandel and Lazear (1992). Punishment may take on many forms, ranging from informal social sanctions to more formalized disciplinary procedures. It also can be represented analytically. In AFG's two-period model, the threat to punish takes the form of a "tit-for-tat" strategy. Each agent works in the first period and then works in the second period only if the other agent worked in the first period. Nikias (2001) provides experimental evidence that under certain conditions, participants play punishment strategies when faced with the AFG incentive scheme.

The principal provides output-based *team incentives* in period 1, meaning that each agent's compensation is increasing in team output and that each agent prefers both agents working to both shirking. Recall that I control for risk preferences by basing experimental payments on the expected values of the stochastic distributions. Thus, the stochastic element of output is removed, and the requirement that each agent's compensation is increasing in team output is equivalent to the requirement that each agent's compensation is increasing in total team effort. Under the horizontal system, if both agents work, then each agent receives 20 points from the principal, resulting in a net profit of 10 points after deducting the 10-point cost of effort. If only one agent works, then each agent receives 14 points, but only one incurs the 10-point cost of effort. If both agents shirk, then each agent receives 8 points. As represented in Exhibit 2, Panel A, this incentive structure is essentially a prisoner's dilemma game, in that each agent's dominant strategy is to shirk even though both could do better by working. Thus, in the one-period (or finitely repeated) version of the game, the unique subgame perfect Nash Equilibrium has both agents shirking. However, the principal can potentially induce the agents to work in period 1 by changing the incentive structure in period 2.

In period 2, summarized in Exhibit 2, Panel B, the principal creates an opportunity for the agents to punish each other. In this period, if both agents work, then each agent receives 20 points. If only one agent works, then each receives 10 points. If both agents shirk, then each receives 0 points. Thus, each agent is indifferent to working or shirking. (The individual agent's gain in compensation from working exactly offsets the 10-point deduction for effort aversion.) However, each agent strongly prefers that the other agent work. Each agent can therefore threaten to punish his/her teammate for shirking in period 1 by shirking in period 2. All four cells of the period 2 subgame represent Nash Equilibria, so either agent can credibly threaten to move from the work-work equilibrium to punish deviant behavior in period 1. Thus, the punishment strategy is a Nash Equilibrium in the two-period game. Further, it is a Nash Equilibrium in the n-period game formed by repeating the two-period game $n/2$ times.[12] Note that the behavioral outcome of this equilibrium is that both agents will work each period. Therefore, if both agents use the punishment strategy, they will each net 20 points over the two-period game. An agent who shirks in period 1 will only net 14 points over the two-period game (14 points in period 1 and 0 points in period 2, when s/he is punished by the other agent).

---

[12] The punishment strategy could be sustained as a Nash Equilibrium in an n-period game by using the individual (punishment phase) incentives only in period n. However, to increase the salience of the opportunity for punishment, I use the individual incentives in all even-numbered periods.

EXHIBIT 2
Horizontal Incentive System

Panel A: Normal Form Representation of the Horizontal Incentive System—Period 1[a]

	Agent 2	
Agent 1	*s*	*w*
*s**	8, 8**	14, 4
w	4, 14	10, 10

Panel B: Normal Form Representation of the Horizontal Incentive System—Period 2[b]

	Agent 2	
Agent 1	*s*	*w*
*s**	0, 0**	10, 0
w	0, 10	10, 10

---

* Represents effort choice (*s* = shirk, *w* = work).

** Represents the payoff to Agent 1 and Agent 2, respectively.

[a] Period 1 of the Horizontal Incentive System is essentially a prisoner's dilemma, in that the unique Nash Equilibrium is for each agent to shirk, but the Pareto optimal outcome is for each agent to work.

[b] In period 2, all four cells represent Nash Equilibria, because each agent is indifferent to his/her own actions. For example, if Agent 1 shirks, Agent 2 earns 0 points regardless of whether s/he shirks or works. Likewise, if Agent 1 works, Agent 2 earns 10 points regardless of whether s/he shirks or works. Therefore, period 2 can be used as a punishment period, which sustains a tit-for-tat strategy in the two-period game. The Pareto optimal equilibrium in the two-period game is for each agent to work in period 1 and then to work in period 2 only if the other agent worked in period 1.

A natural question arises at this point: Why doesn't the principal use the period 2 (individual) incentives in all periods? The principal could use these incentives in all periods and induce working as equilibrium behavior. The reason the principal would not want to do this in a conventional principal-agent setting is related to risk aversion. The individual incentives of period 2 impose greater risk on the agents than do the team incentives of period 1. While I have assumed risk-neutrality for simplification, real-world agents would likely be risk-averse, requiring greater pay under higher-risk incentives. Therefore, the principal would want to use the higher-risk individual incentives only as needed to provide for a punishment opportunity.

While the outcome of each agent playing the tit-for-tat or some other punishment strategy is a Nash Equilibrium, it is not unique. Therefore, it does not rule out the possibility that the agents will play some other equilibrium that is not in the principal's best interest. For example, the undesirable outcome in which both agents shirk each period is also an equilibrium. (If one agent's strategy is to always shirk, then the other agent's best response is to always shirk.) However, the punishment equilibrium Pareto dominates all other Nash Equilibria, and so agents who coordinate their actions are more likely to achieve the punishment equilibrium. What determines whether the agents will coordinate their actions and thus achieve the desirable result of working each period? One answer to this question may be related to the relationship between the two agents, as developed in the next section.

Team Identity

As described earlier, a key difference between the horizontal and vertical systems relates to assumptions about side contracting. The vertical system assumes that team members choose their strategies independently, with no coordination or side contracting between the agents. In contrast, the horizontal incentive system relies on the assumption that the agents will cooperate, coordinating their actions through implicit or explicit side contracts. Therefore, to predict the effectiveness of either system in eliciting high levels of effort, one must understand the likelihood of coordinated behavior among the agents. The analytic models treat the level of coordination as an exogenous factor, and do not delve into the processes through which such coordination emerges. Social psychology research, on the other hand, provides a theory for predicting when team members will coordinate their actions. In this section, I introduce Social Identity Theory, and describe how it can be applied to make specific predictions regarding the effectiveness of vertical and horizontal incentive systems.

Social Identity Theory describes the psychological processes that occur when a person self-identifies as a group member. A social identity results from a *self-categorization* process, through which an individual cognitively groups himself/herself with others, based on perceived similarities (Turner 1987a; Hogg 1987). Tajfel and Turner (1986) describe this process as a transition, through which a person stops thinking like a unique individual and instead begins to think like a representative of a group. As described by Brewer and Schneider (1990, 170), there is "a fundamental depersonalization of the self-concept."

The adoption of a social identity affects how information is interpreted and how decisions are made (Lembke and Wilson 1998). For example, King (2002) demonstrates that a sense of social identity among auditors mitigates their tendency to over-rely on clients' nonbinding communication. In a team setting, the shift from an individual to a team perspective will change team members' beliefs about how their actions affect outcomes (Wech et al. 1998). In a highly identified team, team members are likely to believe that they can influence outcomes through collective versus individual actions. Thus, they will coordinate their actions and focus on joint rather than individual outcomes (Brewer 1979).

Several studies have investigated the effect of social identity on contributions to public goods, finding that such contributions increase when participants categorize themselves into a group with other participants (Brewer and Kramer 1986; De Cremer and van Vugt 1998; Kramer and Brewer 1984). Similarly, Wit and Wilke (1992) report that individuals who self-categorize themselves as group members achieve greater cooperation in several types of social dilemmas.[13] These social dilemma experiments are particularly relevant for the current study, because both the horizontal and vertical incentive systems of interest have features that relate to social dilemmas. Specifically, in a social dilemma, it is individually rational for each person to defect (behave opportunistically), but each person is better off if all choose to cooperate than if all choose to defect (Dawes 1980). The horizontal and vertical incentive systems demonstrate a similar tension between individual rationality and social welfare.[14]

---

[13] The settings for these studies vary in several important ways from the current study. Most notably, while the current study focuses specifically on incentive systems based on mutual monitoring, the prior studies do not allow participants to observe each other's actions or to communicate with each other. Further, prior investigations of social identity in social dilemma settings have restricted interaction, including opportunities for communication. A team setting generally creates significant opportunities for communication and, thus, in the current study, participants are allowed to communicate in writing.

[14] The horizontal incentive system is not a social dilemma, per se. As described earlier, a social dilemma is characterized by a situation where the unique Nash Equilibrium is not Pareto optimal. This is not the case in the horizontal system, as the outcome that has each person playing a "tit-for-tat" strategy is both a Nash Equilibrium and a Pareto optimal outcome. However, this solution depends entirely on the multiperiod nature of the game, with "enforcement periods" used to sustain cooperation in earlier periods. These earlier periods, standing alone, are essentially risky prisoner's dilemmas. The enforcement periods create an interesting twist.

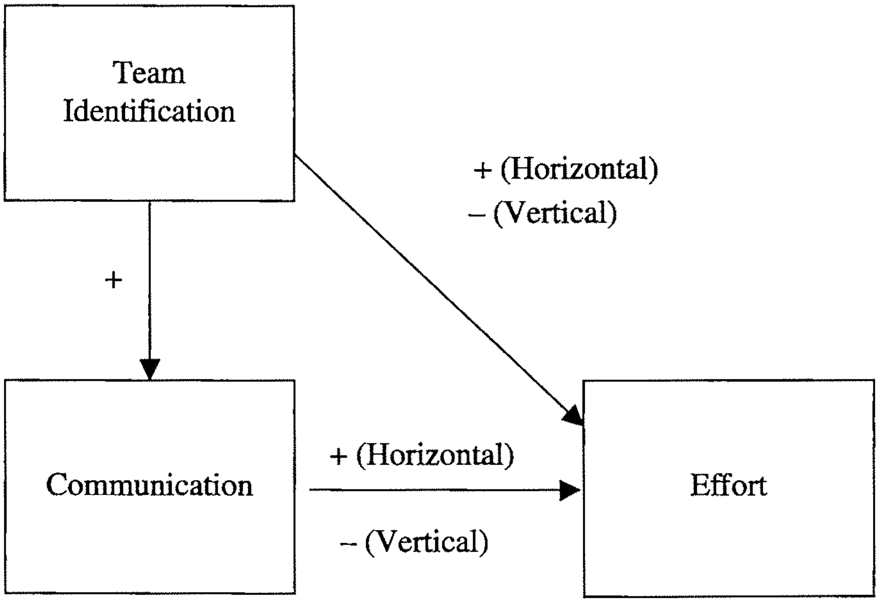
(*continued on next page*)

Hypotheses

This study's primary prediction is an interaction between the incentive system and the level of team identity, such that the effectiveness of an incentive system is either enhanced or degraded by a strong team identity. This prediction is hypothesized formally in H1, following a description of the process resulting in this interaction. Figure 1 illustrates the process.

The reasoning depends on a self-categorization process, by which members of highly identified teams cognitively group themselves with their teammates. As a result, the team, rather than the individual, becomes the primary cognitive unit of analysis. In considering strategy choices, the team members become more attuned to the interrelatedness of their actions, focusing on the ways in which they can jointly affect outcomes. This change in cognitive focus will lead them to choose strategies that are mutually beneficial (or cooperative) in nature. Thus, the cognitive change that defines team identity will have a *direct effect* on the level of cooperation achieved. In other words, highly identified teams are more likely than other teams to reach Pareto optimal outcomes. Team identity also operates through an *indirect effect*, resulting from the efforts of highly identified teams to actively coordinate their strategies. Coordination is enhanced by communication (Cooper et al.

FIGURE 1
The Hypothesized Effects of Team Identity on Effort
(Direct and Indirect Effects)



Footnote 14, continued
While these enforcement periods serve the specific purpose of increasing cooperative behavior, by creating opportunities for retribution, there is some possibility that their usefulness will be undermined by social identity. That is, if a strong team identity reduces participants' willingness to punish one another for not cooperating, it may have the indirect effect of reducing cooperative behavior, a result opposite of that found in prior research on social identity in social dilemmas.

1992), so it follows that highly identified teams will engage in greater communication, leading to an increased probability of arriving at cooperative (Pareto optimal) outcomes.

Thus, I predict that team identity, operating through both direct (cognitive) and indirect (communication-mediated) effects, will lead to greater cooperation. However, cooperation means different things under the two incentive systems. Recall that under the vertical incentive system, pay is based on reported effort. This opens the possibility for collusion between the two agents—they can each shirk and report that the other worked. This collusive outcome is not a Nash Equilibrium, because each agent has strict incentives to report truthfully. However, the collusive outcome is Pareto superior (from the agents' point of view) to the principal's preferred outcome of working and truthful reporting. Under the vertical incentive system, therefore, agents who collude are likely to choose *lower* effort levels than those who choose their strategies independently.

The horizontal incentive system, on the other hand, bases pay on team output and offers no opportunities for collusion. Indeed, under this system, the principal assumes that the agents will side contract to the principal's advantage. There are a number of equilibria (e.g., both agents shirk each period) in the multiple-period game, but the cooperative (Pareto optimal) outcome is the one preferred by the principal. In this equilibrium, each agent plays a punishment strategy, with the result that each agent works each period. Under the horizontal incentive system, therefore, agents who cooperate are likely to choose *higher* effort levels than those who choose their strategies independently.

In summary, team identity should lead to greater cooperation, which can have either a positive or negative effect on effort, depending on the incentive system in place. This is the primary hypothesis, stated in the alternate form.

> **H1:** The incentive system will interact with team identity, such that the effectiveness (in terms of the level of effort elicited) of the horizontal incentive system will be enhanced by a strong team identity, while the effectiveness of the vertical incentive system will be degraded by a strong team identity.

The second hypothesis focuses on the *process* resulting in the predicted interaction between team identity and the incentive system. Specifically, H2 tests the two paths through which team identity is expected to influence the effectiveness of the incentive system—the direct, or cognitive, path, and the indirect, or communication-mediated, path.

> **H2:** Team identity will have both a direct and an indirect, communication-mediated, effect on the effectiveness of the incentive system.

In the discussion to this point, the incentive system and the level of team identity have been considered independent and exogenous factors, reflecting the experimental manipulations. However, *after repeated interaction*, an endogenous relationship can arise between these two constructs. Specifically, the type of incentive system is likely to affect the level of identity that a team achieves, because the two incentive systems emphasize different aspects of the agents' relationship. The horizontal system is geared toward a team mentality. That is, the principal remains unaware of individual contributions and each person's compensation is based on the team's total output. The vertical system, on the other hand, eschews the cooperative spirit of a team. Instead, team members tattle on one another. Individual compensation is based on peer reports, and team members receive bonuses for making unfavorable reports on their peers. For these reasons, the type of incentive system is expected to affect the team members' feelings about one another, and thus the level of team identity. This leads to the final hypothesis.

> **H3:** After repeated interaction, individuals compensated according to the horizontal incentive system will experience a higher level of team identity than will individuals compensated according to the vertical incentive system.

III. METHOD

This experiment uses a 2 × 2 (Incentive System × Team Identity) between-subjects factorial design. Sixteen experimental sessions each involved eight participants recruited from graduate business classes at a large university (i.e., 128 participants in total). As participants arrived, they were randomly assigned to two groups of four, with each group identified by a different color. For the instructional phase, each participant was seated with members of his/her color group.

Color groups were used to manipulate team identity. The wording of the instructions, along with seating assigned by color and the use of colored props, increased the salience of the color groups, facilitating the participants' self-categorizations into these groups. The presence of two color groups in each session promoted this process, as prior research has shown that a team identity may be clarified by in-group/out-group comparisons (Abrams and Hogg 1990). Teams of two were required for the actual experimental task. In the high-identity condition, each team was formed by pairing two members from the *same* color group, whereas in the low-identity condition, each team was formed by pairing two members from *different* color groups.[15] For a real-world analogue to this manipulation, suppose that the color groups represent functional roles (accounting, engineering, etc.). The low-identity teams then represent cross-functional teams, while the high-identity teams represent uni-functional teams.

While seated with their color groups, participants read a scenario, in which they were instructed to assume the roles of two division managers making effort (or more generally, resource allocation) decisions. The scenario described the incentive system, manipulated to represent the two systems of interest. Participants were informed that they would be compensated in cash at a rate of 10 points per dollar, based on the decisions made and the incentive system in place. Because risk preferences were not relevant to this study, the stochastic element was eliminated, and each incentive system was presented in expected value terms.

After the instructions were read and reviewed, the members of each team were seated at opposite ends of a table, where they completed the experimental task. The task was for each participant to choose the level of resources that his/her division would provide. Participants were asked to choose high or low resources, rather than to choose work or shirk, because of the concern that the latter labels might invoke a value judgment and response.[16] The two members of each team made these decisions simultaneously, circling high or low on Resource Decision Forms. The experimenter then showed each participant the form completed by his/her teammate, capturing the notion of mutual monitoring. Participants were informed that in addition to circling "high" or "low" on their Resource Decision

---

[15] While it may seem improbable that *ad hoc* assignments to color groups would engender a sense of group identity, psychology studies on this "minimal group paradigm" are quite robust. Favoritism for in-groups over out-groups, even for explicitly random groups, has been demonstrated in numerous studies. For a review, see Turner (1987b).

[16] From this point forward, the terms "work" and "shirk" are used interchangeably with choosing high and low resources.

Forms, they could write whatever else they would like.[17] Because these forms were shown to their teammates, this allowed for intra-team communication. In the vertical incentive system, each participant also submitted a Report Form, revealing (either truthfully or not) the level of effort selected by his/her teammate. Each session consisted of 20 periods, following which participants completed a post-experimental questionnaire of demographic and process-related questions.[18]

IV. RESULTS

Hypothesis 1

Hypothesis 1 predicts an interaction between the level of team identity and the incentive system.[19] The dependent variable is the effectiveness of the incentive system. This construct is operationalized as "total effort," the cumulative number of times team members chose to work over the course of the experiment. Because under the horizontal incentive system, the even periods are essentially enforcement periods, aimed at eliciting high effort in odd periods, this analysis considers only odd periods.[20] Therefore, this variable can take on any value between 0 and 20, with 20 representing a team for which each member chooses to work all 10 odd periods. Table 1 presents descriptive statistics, and Table 2 presents the results of the 2 × 2 ANOVA.

The hypothesized interaction between team identity and the incentive system is highly significant ($F = 9.38$, $p < 0.01$). Furthermore, simple effects analysis (in Panel B of Table 2) reveals that under the vertical incentive system, high identity teams chose *lower* levels of effort than low identity teams ($F = 3.52$, $p < 0.04$ one-tailed), but that under the horizontal incentive system, high identity teams chose *higher* levels of effort than low identity teams ($F = 6.03$, $p < 0.01$ one-tailed). Thus, the effectiveness of the vertical incentive system is degraded by a strong team identity, while the effectiveness of the horizontal incentive system is enhanced by a strong team identity, supporting H1. This result provides a possible explanation for the observation that the horizontal approach is more common in practice than the vertical approach. Firms using teams often do so to capture gains from cooperation, information-sharing, and productive synergies. A number of consulting firms (e. g., Teambuilding, Inc. and Team Builders Plus) specialize in training aimed at creating a sense of "oneness" among team members. Paradoxically, the results presented here show that the team spirit many companies try to instill actually undermines the incentive effects of the vertical approach. Thus, these firms are likely to design their incentive systems using a horizontal approach, which becomes more effective in the presence of a strong team identity.

Supplemental Analysis on Hypothesis 1

While the primary analysis for H1 focuses on the effort levels, supplemental analysis can provide insight into the social processes through which these effort levels were building may provide greater benefits to the firm than an emphasis on the team members' control over such things as compensation, promotions, and workload distribution.

---

[17] This instruction was provided verbally. The administrator was unaware of the experimental condition at the time of providing this instruction in order to eliminate the potential for experimenter bias.

[18] For the vertical system, this entailed 20 replications of the one-period game, while for the horizontal system, it entailed 10 replications of the two-period game.

[19] A post-experimental question served as a manipulation check for the team identity variable. When asked what their feelings had been at the beginning of the experiment, participants in the high-identity condition reported that they considered the person with whom they were paired to be a "teammate" to a greater extent than did participants in the low-identity condition ($F = 7.69$, $p < 0.01$).

[20] If all periods are used for both systems, then the results are inferentially identical. Further, if only odd periods are used for the horizontal system while all periods (divided by 2 for scaling) are used for the vertical system, then the results are inferentially identical. Finally, results are inferentially identical for analysis using only the first ten periods or only the last ten periods.

Hypothesis 2

Hypothesis 2 is aimed at unraveling the process by which team identity influences the effectiveness of incentive systems, suggesting that the relationship occurs through both direct (cognitive) and indirect (communication) effects. This model is tested using path analysis.[22] Under this structural equations approach, a single covariance matrix is created and used to simultaneously estimate all the links in the model (Kline 1998).[23] The primary measure of fit is a Chi-squared statistic, which tests the null hypothesis that the proposed model is a good fit for the data. This statistic is not statistically significant ($\chi^2 = 0.15$, $p > 0.69$), indicating that the model is a good fit.[24]

The independent and dependent variables for this analysis are the same as those used in the ANOVA for H1. However, an additional mediating variable, the level of communication, is added to the analysis. Recall that participants were allowed to share written communication with their teammates. The level of communication is captured by counting the number of words written by each team.[25]

The standardized path coefficients for this model are presented in Figure 2. The model confirms a significant direct relationship between team identity and effort. That is, there is a team identity × incentive system interaction ($\chi^2 = 6.06$, $p < 0.02$). Note that this result is fundamentally different from that presented for H1. While H1 tested for an interaction between team identity and the incentive system, that test was aimed at identifying only the total effect, not its direct and indirect components. The path analysis is aimed at unraveling these two components. Thus, the significant interaction here suggests a moderating influence of the incentive system on the *direct* effect of team identity on effort. Under the horizontal incentive system, a strong team identity increases the level of effort ($t = 2.80$, $p < 0.01$ one-tailed), whereas under the vertical incentive system, the direct effect of team identity is not statistically significant ($t = -1.08$, $p < 0.15$ one-tailed). However, there is weak evidence of a direct effect, even under the vertical system. Recall that the level of effort is assessed for odd periods only. If all 20 periods of data are used instead, then this direct effect is marginally significant ($t = -1.30$, $p < 0.10$ one-tailed).

I now address the indirect (communication) effect of team identity on effort. As hypothesized, high team identity leads to significant increases in communication ($t = 2.55$, $p < 0.01$ one-tailed). There is also a marginal communication × incentive system interaction ($\chi^2 = 3.16$, $p < 0.08$), meaning that the effect of communication on effort depends on the incentive system in place. Under the vertical incentive system, communication is negatively associated with effort ($t = -1.82$, $p < 0.04$ one-tailed), suggesting that team members used communication to collude with one another, leading them to reduce their support an indirect (communication-mediated) effect of team identity on effort for both the horizontal and vertical incentive systems. However, it should be noted that structural equations-based path analysis cannot be legitimately performed on a dichotomous measure. Therefore, the analysis using the dichotomous measure does not simultaneously estimate all the paths of the model, and should be interpreted with caution.

---

[22] Path analysis is the general term for structural equations modeling when no latent variables are included.

[23] Because of the predicted interactions, the simpler method of mediation analysis proposed by Baron and Kenny (1986) is not appropriate. Instead, I use a multisample approach with nested model comparisons. (For a detailed explanation of this procedure, see Rigdon et al. 1998.)

[24] This evidence is corroborated by the Tucker-Lewis Index, also known as the Non-Normed Fit Index. This value indicates the proportion of improvement of the model over a null model. The value of 1.30, indicating a 130 percent increase over the null model, is well above the generally accepted cut-off value of 90 percent (Kline 1998, 131).

[25] While I did not conduct a formal content analysis of the communication transcripts, casual observation of the content suggests no differences among the experimental conditions. The vast majority of communication was aimed at problem solving and coordination.

In summary, the model predicted in H2 is fully supported under the vertical incentive system. Specifically, under the vertical incentive system, the effort-reducing effect of team identity has both a direct and an indirect (communication-mediated) component. Under the horizontal incentive system, the effort increasing effect of team identity also appears to have direct and indirect components. Under this condition, while high identity did lead to increased communication, the effect of communication on effort was driven by the mere presence of communication and not by the amount of communication.

Hypothesis 3

While the analysis to this point views team identity and the incentive system as two exogenous and independent constructs, H3 predicts that *after repeated interaction,* individuals compensated according to the horizontal incentive system will experience a higher level of team identity than will individuals compensated according to the vertical incentive system (i.e., team identity has an endogenous component in addition to its exogenous manipulation). The dependent variable for this test is ending team identity, measured through the post-experimental survey. Each participant was asked, "At the end of the session, to what extent did you perceive yourself to be a teammate of the person you were paired with?" The Likert scale response to this question is averaged for each pair.

Table 3 reports a two-way (team identity × incentive system) ANOVA. Team identity (representing the *manipulated* high- or low-identity condition) is included because this manipulation is expected to have a lasting effect. That is, a team that begins the process with a high team identity is likely to maintain a relatively high level of identity, regardless of the incentive system in place. Indeed this is the case, as teams in the high-identity

TABLE 3
H3 Results—ANOVA on Ending Team Identity[a]

Factor	dof	Mean Square	F	p-value
Team Identity (as manipulated)[b]	1	17.54	6.23	<0.01*
Incentive System[c]	1	5.94	2.11	<0.08*
Incentive System × Team Identity	1	4.79	1.70	0.20**
Error	60	2.82		

* The p-values for the team identity and incentive system main effects are reported on a one-tailed basis, given the directional predictions for these effects.

** The p-value for the incentive system × team identity interaction is reported on a two-tailed basis, because no directional prediction is made for this effect.

[a] Ending Team Identity is measured via the Likert-scale response to the post-experimental question, "At the end of the session, to what extent did you perceive yourself to be a teammate of the person you were paired with?" The responses are averaged for each pair, and so each team is considered one independent observation.

[b] Team Identity is manipulated between subjects at two levels. High-identity teams were formed from two members of the same color group. Low-identity teams were formed from one member of each color group. The main effect observed indicates that the manipulation of team identity was still in effect at the end of the experiment.

[c] Incentive System is manipulated consistent with the two incentive schemes being studied. Participants accumulated points based on the incentive system formulas. The points were converted to cash at the conclusion of the session.

condition self-reported higher ending levels of team identity than teams in the low-identity condition ($F = 6.23$, $p < 0.01$ one-tailed). This result is essentially a manipulation check. A marginally significant main effect for the incentive system ($F = 2.11$, $p < 0.08$ one-tailed) also exists, suggesting that after repeated interaction, teams operating under the horizontal incentive system experienced higher levels of team identity than did teams operating under the vertical incentive system. Therefore, H3 is marginally supported.

While the incentive system × team identity interaction term for this ANOVA is not significant ($F = 1.70$, $p > 0.19$), it is nonetheless important to point out that the incentive system main effect is driven by the vertical system.[26] The small difference in ending team identity between the two systems in the high-identity condition (6.06—horizontal, 6.00—vertical) is likely due to a ceiling effect. That is, even in the low-identity condition, participants under the horizontal system reported relatively high levels of ending team identity (5.56 out of a possible 7.00). Thus, there is little room for an incremental level of ending team identity in the high identity condition.

V. SUMMARY AND CONCLUSIONS

This study presents evidence that the effectiveness of a financial incentive system can be enhanced or degraded by a sense of team identity, depending on specific characteristics of the incentive system. A strong team identity increases the level of coordination among the agents. This degrades the effectiveness of a vertical incentive system, in which the agents report observations of their teammates' efforts to the principal. Coordination undermines the incentive system, as individuals collude against the principal, choosing low levels of effort and reporting falsely. Conversely, the effectiveness of a horizontal incentive system, based on peer-enacted control, is enhanced by a strong team identity. Because no collusive opportunities exist under this incentive system, a strong team identity serves to help teams to reach a cooperative solution, as desired by the principal. The evidence presented here suggests that under both incentive systems, the effect of team identity on cooperation has both direct (based on a change of cognitive focus) and indirect (communication-mediated) components.

Finally, evidence suggests that team identity is partially endogenous, in that it is affected by the incentive system in place. After repeated interaction, the horizontal system, with its focus on team outcomes rather than individual contributions, leads to higher levels of team identity than does the vertical incentive system, where individuals act as whistle-blowers for the principal.

This research provides useful insights for both theory and practice. From a practical standpoint, this study provides evidence on the usefulness of mutual monitoring as an alternative to traditional methods of measuring individual production inputs. In many joint production settings, workers hold high-quality information on coworkers' inputs, and firms can benefit by understanding how to use this knowledge base to make informed decisions in such areas as resource allocation, compensation, and promotion. The results of this study suggest that when the team has achieved a high level of identity, the most effective way to use this information is likely horizontal in nature, delegating responsibility for control to self-managed teams, rather than extracting the information through reporting mechanisms.

---

[26] Simple effects analysis demonstrates that in the low-identity condition, ending team identity was significantly higher in the horizontal incentive system than in the vertical incentive system ($F = 3.79$, $p < 0.03$ one-tailed). In the high-identity condition, the level of ending team identity was not significantly different under the two incentive systems ($F = 0.01$, $p < 0.93$).

This study thus helps explain why firms have more readily embraced horizontal incentive systems than vertical incentive systems.

Further, this study provides evidence on the types of incentive systems that are most likely to be effective for different types of teams. For example, Scott (1997) demonstrates that permanent teams tend to develop higher levels of identity than do part-time or temporary teams. One inference from the results presented here is that the horizontal approach is particularly appropriate for work teams involved in manufacturing or other permanent production-related tasks. Thus, by focusing on the consequences of team identification, this study aids in understanding the causal linkages among team characteristics, incentives, and performance.

From a theoretical standpoint, this study draws from both economics and psychology, demonstrating that concepts from psychology may be useful in understanding the economic incentives addressed by agency theory. The study therefore adds to an expanding body of literature examining variables that interact with financial incentives to influence performance (Bonner et al. 2000; Bonner and Sprinkle 2002; Drake et al. 1999; Sprinkle 2000). The paper contributes more specifically to the relatively undeveloped literature on the effect of incentives in a team environment (Drake et al. 1999; Fisher et al. 2002; Rankin and Sayre 2000; Rankin 2003).

Certain factors limit the study's generalizability. For example, one may view these results as implying that firms should *always* use the horizontal system, because it performs better than the vertical incentive system when the team identity is high and no worse when the team identity is low. However, this study was not designed to directly compare these two systems, and any such comparisons should be interpreted with caution, due to the study's suppression of risk preferences. That is, to simplify the laboratory experiment, the stochastic element was removed, and the incentive systems were presented to participants in expected value terms. Such an abstraction is appropriate, because risk preferences are not relevant to the hypotheses investigated in this study. However, risk preferences are an outside factor that may affect the relative desirability of the two systems.

I also make the simplifying assumption that effort is discrete and perfectly observable, an attribute almost certainly not characteristic of real-world settings. While team members may observe the amount of time a coworker spends on a particular task, it is considerably more difficult to observe the intensity or creativity with which s/he works.[27] This study compares two analytic approaches to incentive contracting, and these simplifying assumptions help to enhance comparability of results across the two incentive systems. However, it is unclear how the results would change if effort were imperfectly observable. For example, in their experimental investigation of several budget-based group incentive plans, Sprinkle et al. (2003) demonstrate different effects on effort duration, which is easily observed, and effort intensity, which is less easily observed. This limitation offers an opportunity for future research, and a body of psychology literature may prove helpful. Social Comparison Theory (Kunda 1999, 494) provides a rich theory for understanding the more subtle aspects of mutual monitoring. Thus, future work in this area may follow the approach used here of melding insights from both economics and psychology to understand complex managerial accounting issues.

---

[27] These dimensions of effort are not only unobservable, but are also likely immeasurable. This leads to difficulty in meeting another assumption—common knowledge of the probability distribution of outcomes given effort.

APPENDIX
TECHNICAL DETAILS

In this appendix, I introduce a model of team production, including the opportunity for mutual monitoring. I then demonstrate how each type of incentive system (vertical versus horizontal) may be used in this setting. Consider two risk-neutral and effort-averse agents (a and b) engaged in joint production over n periods. Each agent i selects an effort level $e_i \in \{shirk, work\}$ and output $x \in \{high, low\}$ is a function of these effort choices. Specifically, $x = f(e_a, e_b, \theta)$, where θ is a random variable representing uncertainty in the production process. As reflected in θ, output is stochastically related to the agents' effort levels, resulting in the conditional probability distribution depicted in Exhibit 3. For the remainder of this appendix, I suppress θ and focus on the resultant conditional probability distribution, $p(x_j| e_a, e_b)$, the probability that output is j, given the levels of effort selected by the two agents.

Each agent i maximizes expected utility $u_i = g(e_i, m_i)$, which is increasing in monetary wages (m_i) and decreasing in effort (e_i). This function is additively separable in its two operands. As is common in many experimental studies, I make the simplifying assumption that utility can be measured in dollars. Therefore, the utility function may be represented as $u_i = m_i - c(e_i)$, where $c(e_i)$ is a function converting effort levels to costs. Assume that the principal prefers each agent to work (w) rather than shirk (s).[28] However, $c(w) > c(s)$ due to effort aversion, so *ceteris paribus*, the agents will both prefer shirking. Next, I will explore the use of two types of incentive systems to induce the agents to work in this setting.

Vertical Incentive System

Under the vertical system, agent i's wages m_i have two additive components—effort pay and reporting pay. Formally, $m_i = E_i + R_i$, where:

$$E_i \text{ (effort pay)} = \begin{cases} \hat{u}_i + c(w) & \text{if } r_i = \text{work} \\ \hat{u}_i + c(s) - \delta & \text{if } r_i = \text{shirk} \end{cases} \tag{1}$$

$$R_i \text{ (reporting pay)} = \begin{cases} \beta & \text{if } r_j = \text{shirk and } e_j = \text{shirk} \\ \gamma & \text{if } r_j = \text{shirk and } e_j = \text{work} \\ 0 & \text{otherwise (i.e., if } r_j = \text{work)} \end{cases} \tag{2}$$

$r_i \in \{shirk, work\}$ is agent i's effort, as reported by agent j
$\hat{u}_r$ = agent i's reservation utility
$\delta \geq \beta > 0;$[29] $\gamma < 0;\ i \neq j$

Intuitively, effort pay represents wages for effort as reported by the other agent, and each agent will be paid his/her reservation utility (plus an amount equal to the disutility of effort) if the report indicates that s/he has worked.[30] If the report indicates that s/he

---

[28] I treat the principal and the incentive system exogenously, with all analysis directed to the agents' subgame.

[29] The requirement that $\delta \geq \beta$ is not a technical requirement, but rather a practical one. If $\beta > \delta$, then a collusive opportunity exists for the agents to increase their payoffs by agreeing to shirk and report truthfully. This collusive outcome (not a Nash Equilibrium) is easily avoided by adding this constraint.

[30] In Ma's original model, the incentive scheme is asymmetric, with one agent reporting and the other agent verifying the first agent's report. Further, Ma assumes that each unique pair of actions by the two agents leads to a unique probability distribution of outputs, an assumption not met in the simplified setting used in this experiment.

EXHIBIT 3
Conditional Probability Distribution

Effort Levels	**Probabilities of Output Levels** Low	**Probabilities of Output Levels** High
Both Shirk	75%	25%
Agent 1 Shirks/Agent 2 Works	50%	50%
Agent 1 Works/Agent 2 Shirks	50%	50%
Both Work	25%	75%

Three assumptions are made regarding this conditional distribution:

1) First-order stochastic dominance.
2) Non-moving support (the stipulation that the support of x (output) is the same for any combination of efforts by the two agents (Holmstrom 1979)).
3) Symmetry—this assumption is not required, but is made for simplicity.

shirked, then a penalty of δ is subtracted. Reporting pay is the wage component that induces truthful reporting, because it pays a bonus (penalty) to the agent who accuses his/her teammate of shirking, if that report is truthful (false). In this study, the following specific parameters are used, resulting in the normal form game tree found in Exhibit 1.

$$\hat{u}_i = \$10.00^{31}$$

$$c(s) = \$0.00, \qquad c(w) = \$10.00$$

$$\beta = \$10.00$$

$$\delta = \$10.00$$

$$\gamma = -\$20.00.$$

This game can be solved by backward induction, resulting in the unique subgame perfect Nash Equilibrium in which each agent works and each agent reports truthfully. Further, if this game is repeated finitely, the unique subgame perfect Nash Equilibrium in the n-period game will have the agents working and reporting truthfully in each period (Gibbons 1992, 84).

Horizontal Incentive System

As detailed in Exhibit 2, Panel A, the principal provides output-based *team incentives* in period I, meaning that each agent's compensation is increasing in team output and that each agent prefers both agents working to both shirking. Precisely,

$$m_i = \begin{cases} \hat{u}_i + c(w) + \lambda_1 & \text{if } x = \text{high} \\ \hat{u}_i + c(w) + \lambda_1 - \dfrac{\lambda_1}{p(low|w,w)} & \text{if } x = \text{low} \end{cases} \tag{3}$$

---

[31] It is not necessary that the participants' reservation utilities exactly equal \$10.00. In the laboratory, it is only important that the reward structure achieve Smith's (1982) dominance precept, meaning that it must dominate any subjective, nonmonetary costs or values that the subjects hold for participating in the experiment.

Arya, A., and J. Glover. 1996a. The role of budgeting in eliminating tacit collusion. *Review of Accounting Studies* 1 (3): 191–205.

———, and ———. 1996b. Verification of historical cost reports. *The Accounting Review* 71 (2): 255–259.

———, J. Fellingham, and J. Glover. 1997. Teams, repeated tasks, and implicit incentives. *Journal of Accounting and Economics* 23 (1): 7–30.

Baiman, S., and M. V. Rajan. 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review* 70 (4): 557–579.

Baron, R. M., and D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6): 1173–1182.

Barron, J. M., and K. P. Gjerde. 1997. Peer pressure in an agency relationship. *Journal of Labor Economics* 15 (2): 234–254.

Bonner, S. E., R. Hastie, G. B. Sprinkle, and S. M. Young. 2000. A review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. *Journal of Management Accounting Research* 12: 19–64.

———, and G. B. Sprinkle. 2002. The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society* 27 (45): 303–345.

Brewer, M. B. 1979. In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin* 86(2): 307–324.

———, and R. M. Kramer. 1986. Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology* 50 (3): 543–549.

———, and S. K. Schneider. 1990. Social identity and social dilemmas: A double-edged sword. In *Social Identity Theory: Constructive and Critical Advances*, edited by D. Abrams and M. A. Hogg. New York, NY: Springer-Verlag New York, Inc.

Coates, D. E. 1998. Don't tie 360 feedback to pay. *Training* 35 (9): 68–78.

Cohen, S. G., and D. E. Bailey. 1997. What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management* 23 (3): 239–290.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross. 1992. Communication in coordination games. *The Quarterly Journal of Economics* 107 (2): 739–771.

Dawes, R. M. 1980. Social dilemmas. *Annual Review of Psychology* 31: 169–193.

De Cremer, D., and M. van Vugt. 1998. Collective identity and cooperation in a public goods dilemma: a matter of trust or self-efficacy? *Current Research in Social Psychology* 3 (1). Available at: http://www.uiowa.edu/~grpproc/crisp/crisp.html.

DeMatteo, J. S., L. T. Eby, and E. Sundstrom. 1998. Team-based rewards: Current empirical evidence and directions for future research. *Research in Organizational Behavior* 20: 141–183.

Demski, J. S., and D. Sappington. 1984. Optimal incentive contracts with multiple agents. *Journal of Economic Theory* 33: 152–171.

———, ———, and P. T. Spiller. 1988. Incentive schemes with multiple agents and bankruptcy constraints. *Journal of Economic Theory* 44: 156–167.

———. 1997. *Managerial Uses of Accounting Information*. Norwell, MA: Kluwer Academic Publishers.

Drake, A. R., S. F. Haka, and S. P. Ravenscroft. 1999. Cost system and incentive structure effects on innovation, efficiency and profitability in teams. *The Accounting Review* 74 (3): 323–345.

Dumaine, B. 1990. Who needs a boss? *Fortune* 121 (10): 52–60.

Edwards, M. R., and A. J. Ewen. 1996. How to manage performance and pay with 360-degree feedback. *Compensation & Benefits Review* 28 (3): 41–46.

Fischer, P. E., and J. S. Hughes. 1997. Mutual monitoring and best agency contracts. *Journal of Institutional and Theoretical Economics* 153: 334–355.

Fisher, J. G., L. A. Maines, S. A. Peffer, and G. B. Sprinkle. 2002. Using budgets for performance evaluation: Effects of resource allocation and horizontal information asymmetry on budget proposals, budget slack, and performance. *The Accounting Review* 77 (4): 847–865.

———, ———, ———, and ———. 2003. An experimental investigation of subjective versus objective performance evaluation. Working paper, Kelley School of Business, Indiana University, Bloomington, IN.

Fishman, C. 1996. Whole foods is all teams. *Fast Company* 2: 103.

Frederickson, J. R. 1992. Relative performance information: The effects of common uncertainty and contract type on agent effort. *The Accounting Review* 67 (4): 647–669.

Gibbons, R., and K. J. Murphy. 1990. Relative performance evaluation for chief executive officers. *Industrial and Labor Relations Review* 43 (3): 30–51.

———. 1992. *Game Theory for Applied Economists.* Princeton, NJ: Princeton University Press.

Gillmor, D. 2000. Cisco's leader pushes his vision. *The Inquirer* (May 28): E3.

Haynes, C. M., and S. J. Kachelmeier. 1998. The effects of accounting contexts on accounting decisions: A synthesis of cognitive and economic perspectives in accounting experimentation. *Journal of Accounting Literature* 17: 97–136.

Hewitt Associates, LLC. 1995. *Salary Increase Report.* Lincolnshire, IL: Hewitt Associates LLC.

Hogg, M. A. 1987. Social identity and group cohesiveness. In *Rediscovering the Social Group: A Self-Categorization Theory,* edited by J. C. Turner. Oxford, U.K.: Basil Blackwell Ltd.

Holmstrom, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 10 (1): 74–91.

———. 1982. Moral hazard in teams. *Bell Journal of Economics* 13 (2): 324–340.

Institute of Management Accountants (IMA). 1999. *Counting More, Counting Less, Transformations in the Management Accounting Profession, The 1999 Practice Analysis.* Montvale, NJ: IMA.

Itoh, H. 1991. Incentives to help in multiagent situations. *Econometrica* 59 (3): 611–636.

———. 1993. Coalitions, incentives, and risk sharing. *Journal of Economic Theory* 60 (2): 410–427.

Jackson, J. H., and M. M. Greller. 1998. Decision elements for using 360° feedback. *Human Resource Planning* 21 (4): 18–28.

Kandel, E., and E. P. Lazear. 1992. Peer pressure and partnerships. *Journal of Political Economy* 100 (4): 801–817.

Kelley, M. 2000. Core values keep airline flying high. *Harvard Business School Bulletin* (December). Available at: http://www.hbsworkingknowledge.hbs.edu.

Kidwell, R. E., Jr., and N. Bennett. 1993. Employee propensity to withhold effort: A conceptual model to intersect three avenues of research. *Academy of Management Review* 18 (3): 429–456.

King, R. 2002. An experimental investigation of self-serving biases in an auditing trust game: The effect of group affiliation. *The Accounting Review* 77 (2): 265–284.

Kline, R. B. 1998. *Principles and Practice of Structural Equation Modeling.* New York, NY: The Guilford Press.

Kramer, R. M., and M. B. Brewer. 1984. Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology* 46 (5): 1044–1057.

Kunda, Z. 1999. *Social Cognition—Making Sense of People.* Cambridge, MA: The MIT Press.

Lazega, E. 2000. Rule enforcement among peers: A lateral control regime. *Organization Studies* 21 (1): 193–214.

Lembke, S., and M. Wilson. 1998. Putting the "team" into teamwork: Alternative theoretical contributions for contemporary management practice. *Human Relations* 51 (7): 927–944.

Ma, C. 1988. Unique implementation of incentive contracts with many agents. *Review of Economic Studies* 55 (4): 555–572.

———, J. Moore, and S. Turnbull. 1988. Stopping agents from "cheating." *Journal of Economic Theory* 46: 355–372.

Milgrom, P., and J. Roberts. 1992. *Economics, Organization, and Management.* Englewood Cliffs, NJ: Prentice Hall.

Mookherjee, D. 1984. Optimal incentive schemes with many agents. *Review of Economic Studies* 51 (3): 433–446.

Moser, D. V. 1998. Using an experimental economics approach in behavioral accounting research. *Behavioral Research in Accounting* 10 (Supplement): 94–110.

Nahavandi, A., and E. Aranda. 1994. Restructuring teams for the reengineered organization. *Academy of Management Review* 8 (4): 58–68.

Nikias, A. 2001. Implicit punishment strategies in a multiperiod team setting. Working paper, The Ohio State University, Columbus, OH.

Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37 (1): 763.

Radner, R. 1986. Repeated partnership games with imperfect monitoring and no discounting. *Review of Economic Studies* 53 (1): 43–57.

Ramakrishnan, R. T. S., and A. V. Thakor. 1991. Cooperation versus competition in agency. *Journal of Law, Economics, and Organization* 7 (2): 248–283.

Rankin, F., and T. L. Sayre. 2000. The effects of performance separability and contract type on agent effort. *Accounting, Organizations and Society* 25 (7): 683–695.

———. 2003. The effect of the mutual observability of information and performance measurement on the allocation of effort: An experimental analysis. Working paper, Washington University, St. Louis, MO.

Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker and G. A. Marcoulides. Mahwah, NJ: Lawrence Erlbaum Associates.

Scott, S. G. 1997. Social identification effects in product and process development teams. *Journal of Engineering Technology Management* 14 (2): 97–127.

Smith, V. L. 1982. Microeconomic systems as an experimental science. *The American Economic Review* 72 (5): 923–955.

Sprinkle, G. B. 2000. The effect of incentive contracts on learning and performance. *The Accounting Review* 75 (3): 299–326.

———, Peffer, S. A., and Fisher, J. G. 2003. Budget-based contracts, budget levels, and group productivity. *Journal of Management Accounting Research* 15 (forthcoming).

Tajfel, H., and J. Turner. 1986. The social identity theory of intergroup behavior. In *The Social Psychology of Intergroup Relations*, 2nd edition, edited by S. Worchel, and W. G. Austin. Chicago, IL: Nelson-Hill Publishers.

Thatcher, M. 1996. Allowing everyone to have their say. *Personnel Management* 2 (6): 28–30.

Tirole, J. 1988. The multi-contract organization. *Canadian Journal of Economics* 21 (3): 459–466.

Turner, J. C. 1987a. A self-categorization theory. In *Rediscovering the Social Group: A Self-Categorization Theory*, edited by J. C. Turner. Oxford, U.K.: Basil Blackwell Ltd.

———. 1987b. Rediscovering the social group. In *Rediscovering the Social Group: A Self-Categorization Theory*, edited by J. C. Turner. Oxford, U.K.: Basil Blackwell Ltd.

Varian, H. R. 1990. Monitoring agents with other agents. *Journal of Institutional and Theoretical Economics* 146: 153–174.

Villadsen, B. 1995. Communication and delegation in collusive agencies. *Journal of Accounting and Economics* 19 (23): 315–344.

Waller, W. S. 1995. Decision-making research in managerial accounting: Return to behavioral-economics foundations. In *Judgment and Decision-Making Research in Accounting and Auditing*, edited by R. H. Ashton, and A. H. Ashton. New York, NY: Cambridge University Press.

———. 2001. Behavioral accounting experiments in market and game settings. In *Experimental Business Research*, edited by R. Zwick, and A. Rapoport. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Wech, B. A., K. W. Mossholder, R. P. Steel, and N. Bennett. 1998. Does work group cohesiveness affect individuals' performance and organizational commitment? A cross-level examination. *Small Group Research* 29 (4): 472–494.

Wellins, R. S., W. C. Byham and G. R. Dixon. 1994. *Inside Teams: How 20 World-Class Organizations are Winning Through Teamwork.* San Francisco, CA: Jossey-Bass.

Wells, S. J. 1999. A new road: Traveling beyond 360-degree evaluation. *HR Magazine* 44 (9): 82–86.

Wit, A. P., and H. A. M. Wilke. 1992. The effect of social categorization on cooperation in three types of social dilemmas. *Journal of Economic Psychology* 13 (1): 135–151.

Editorial Data

The following table contains information about turnaround time for manuscripts (including revisions) on which editorial decisions were made in the 12-month period ended May 31, 2003. Turnaround time is the number of days between the date that the manuscript was received *and the date of the editor's letter to the author(s)*:

	Number of Manuscripts	Cumulative Number	Cumulative Percent
0 ≤ Days ≤ 30	60	60	15.42
31 ≤ Days ≤ 60	143	203	52.19
61 ≤ Days ≤ 90	136	339	87.15
91 ≤ Days ≤ 120	38	377	96.92
121 ≤ Days	12	389	100.00

The mean review time was 57.9 days; the median review time was 57 days.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of *The Chicago Manual of Style* (14th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be typed in 12-point font on one side of $8\frac{1}{2} \times 11''$ good quality paper and be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form*. See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings.

Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix a, b, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says ..."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by *The Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Date of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2nd edition. Homewood, IL: Dow Jones-Irwin.
Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, U.K.: Cambridge University Press.
Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–636.
———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.
Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, The University of Texas at Austin.
Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan or the like) should be submitted. Information that might identify the author(s) must be deleted from the instrument.
3. MS should be submitted via email as Microsoft® Word or PDF files to the *TAR* office at taruw@u.washington.edu. A submission letter, a separate cover page, and the MS containing no author identification should be emailed as separate files. The submission fee should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If unable to pay online, a check in U.S. funds, made payable to the American Accounting Association, may be sent to Professor Terry Shevlin, Senior Editor *TAR*, School of Business, Box 353200, University of Washington, Seattle, WA 98195-3200. The submission fee in U.S. funds is $125.00 for members and $200.00 for nonmembers of the AAA. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting. "An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

AUTHOR INDEX TO VOLUME 78

Author	*Title*	*Page*
ADUT, DAVIT (and CREADY and LOPEZ)	Structuring Charges and CEO Cash Compensation: A Reexamination	169
ALI, ASHIQ (and HWANG and TROMBLEY)	Residual-Income-Based Valuation Predicts Future Stock Returns: Evidence on Mispricing vs. Risk Explanations	377
ALLEN, ROBERT D. (and ELDER)	A Longitudinal Field Investigation of Auditor Risk Assessments and Sample Size Decisions	983
ASHBAUGH, HOLLIS (and LaFOND and MAYHEW)	Do Nonaudit Services Compromise Auditor Independence? Further Evidence	611
BAGINSKI, STEPHEN P. (and WAHLEN)	Residual Income Risk, Intrinsic Values, and Share Prices	327
BALDENIUS, TIM	Delegated Investment Decisions and Private Benefits of Control	909
BEATTY, ANNE (and WEBER)	The Effects of Debt Contracting on Voluntary Accounting Method Change	119
BEDARD, JEAN C. (and JOHNSTONE)	Risk Management in Client Acceptance Decisions	1003
BHATTACHARYA, UPTAL (and DAOUK and WELKER)	The World Price of Earnings Opacity	641
BONNER, SARAH E. (and WALTHER and YOUNG)	Sophistication-Related Differences in Investors' Models of the Relative Accuracy of Analysts' Forecast Revisions	679
BOSCHEN, JOHN F. (and DURU and GORDON and SMITH)	Accounting and Stock Price Performance in Dynamic CEO Compensation Arrangements	143
CARCELLO, JOSEPH V. (and NEAL)	Audit Committee Characteristics and Auditor Dismissals following "New" Going-Concern Reports	95

CHEN, PETER F. (and ZHANG)	Heterogeneous Investment Opportunities in Multiple-Segment Firms and the Incremental Value Relevance of Segment Accounting Data	397
CHUNG, HYEESOO (and KALLAPUR)	Client Importance, Nonaudit Services, and Abnormal Accruals	931
CLEMENT, MICHAEL B. (and TSE)	Do Investors Respond to Analysts' Forecast Revisions as if Forecast Accuracy Is All That Matters?	227
CORE, JOHN E. (and GUAY and VERRECCHIA)	Price versus Non-Price Performance Measures in Optimal CEO Compensation Contracts	957
CREADY, WILLIAM H. (and ADUT and LOPEZ)	Structuring Charges and CEO Cash Compensation: A Reexamination	169
DAOUK, HAZEM (and BHATTACHARYA and WELKER)	The World Price of Earnings Opacity	641
DAVIS, JON S. (and HECHT and PERKINS)	Social Behaviors, Enforcement, and Tax Compliance Dynamics	39
DURU, AUGUSTINE (and BOSCHEN and GORDON and SMITH)	Accounting and Stock Price Performance in Dynamic CEO Compensation Arrangements	143
DUTTA, SUNIL	Capital Budgeting and Managerial Compensation: Incentive and Retention Effects	71
EAMES, MICHAEL J. (and GLOVER)	Earnings Predictability and the Direction of Analysts' Earnings Forecast Errors	707
ELDER, RANDAL J. (and ALLEN)	A Longitudinal Field Investigation of Auditor Risk Assessments and Sample Size Decisions	983
FAIRFIELD, PATRICIA M. (and WHISENANT and YOHN)	Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing	353
FERGUSON, ANDREW (and FRANCIS and STOKES)	The Effects of Firm-Wide and Office-Level Industry Expertise on Audit Pricing	429

FRANCIS, JERE R. (and FERGUSON and STOKES)	The Effects of Firm-Wide and Office-Level Industry Expertise on Audit Pricing	429
GLEASON, CRISTI A. (and LEE)	Analyst Forecast Revisions and Market Price Discovery	193
GLOVER, STEVEN M. (and EAMES)	Earnings Predictability and the Direction of Analysts' Earnings Forecast Errors	707
GORDON, LAWRENCE A. (and BOSCHEN and DURU and SMITH)	Accounting and Stock Price Performance in Dynamic CEO Compensation Arrangements	143
GUAY, WAYNE R. (and CORE and VERRECCHIA)	Price versus Non-Price Performance Measures in Optimal CEO Compensation Contracts	957
HECHT, GARY (and DAVIS and PERKINS)	Social Behaviors, Enforcement, and Tax Compliance Dynamics	39
HEFLIN, FRANK (and SUBRAMANYAM and ZHANG)	Regulation FD and the Financial Information Environment: Early Evidence	1
HODDER, LESLIE (and McANALLY and WEAVER)	The Influence of Tax and Nontax Factors on Banks' Choice of Organizational Form	297
HWANG, LEE-SEOK (and ALI and TROMBLEY)	Residual-Income-Based Valuation Predicts Future Stock Returns: Evidence on Mispricing vs. Risk Explanations	377
ITTNER, CHRISTOPHER D. (and LARCKER and MEYER)	Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard	725
JOHNSTONE, KARLA M. (and BEDARD)	Risk Management in Client Acceptance Decisions	1003
JORGENSEN, BJORN N. (and KIRSCHENHEITER)	Discretionary Risk Disclosures	449
JUAREZ-VALDES, LUIS FELIPE (and SWANSON and REES)	The Contribution of Fundamental Analysis after a Currency Devaluation	875

KADOUS, KATHRYN (and KENNEDY and PEECHER)	The Effect of Quality Assessment and Directional Goal Commitment on Auditors' Assessment of Client-Preferred Accounting Methods	759
KALLAPUR, SANJAY (and BALDENIUS)	Client Importance, Nonaudit Services, and Abnormal Accruals	931
KENNEDY, S. JANE (and KADOUS and PEECHER)	The Effect of Quality Assessment and Directional Goal Commitment on Auditors' Assessment of Client-Preferred Accounting Methods	759
KIRSCHENHEITER, MICHAEL T. (and JORGENSEN)	Discretionary Risk Disclosures	449
KUMAR, KRISHNA R. (and VISVANATHAN)	The Information Content of the Deferred Tax Valuation Allowance	471
LaFOND, RYAN (and ASHBAUGH and MAYHEW)	Do Nonaudit Services Compromise Auditor Independence? Further Evidence	611
LARCKER, DAVID F. (and ITTNER and MEYER)	Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard	725
LEE, CHARLES M. C. (and GLEASON)	Analyst Forecast Revisions and Market Price Discovery	193
LOPEZ, THOMAS J. (and ADUT and CREADY)	Structuring Charges and CEO Cash Compensation: A Reexamination	169
LOUIS, HENOCK	The Value Relevance of the Foreign Translation Adjustment	1027
MANRY, DAVID (and TIRAS and WHEATLEY)	The Influence of Interim Auditor Reviews on the Association of Returns with Earnings	251
MAYHEW, BRIAN W. (and ASHBAUGH and LaFOND)	Do Nonaudit Services Compromise Auditor Independence? Further Evidence	611
McANALLY, MARY LEA (and HODDER and WEAVER)	The Influence of Tax and Nontax Factors on Banks' Choice of Organizational Form	297

MEYER, MARSHALL W. (and ITTNER and LARCKER)	Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard	725
MYERS, JAMES N. (and MYERS and OMER)	Exploring the Term of the Auditor-Client Relationship and the Quality of Earnings: A Case for Mandatory Auditor Rotation?....................	779
MYERS, LINDA A. (and MYERS and OMER)	Exploring the Term of the Auditor-Client Relationship and the Quality of Earnings: A Case for Mandatory Auditor Rotation?....................	779
NEAL, TERRY L. (and CARCELLO)	Audit Committee Characteristics and Auditor Dismissals following "New" Going-Concern Reports..	95
NG, TERENCE BU-PEOW (and TAN)	Effects of Authoritative Guidance Availability and Audit Committee Effectiveness on Auditors' Judgments in an Auditor-Client Negotiation Context..	801
OMER, THOMAS C. (and MYERS and MYERS)	Exploring the Term of the Auditor-Client Relationship and the Quality of Earnings: A Case for Mandatory Auditor Rotation?....................	779
PATTERSON, EVELYN R. (and SMITH)	Materiality Uncertainty and Earnings Misstatement ..	819
PAYNE, JEFF L. (and THOMAS)	The Implications of Using Stock-Split Adjusted I/B/E/S Data in Empirical Research...............	1049
PEECHER, MARK E. (and KADOUS and KENNEDY)	The Effect of Quality Assessment and Directional Goal Commitment on Auditors' Assessment of Client-Preferred Accounting Methods...............	759
PERKINS, JON D. (and DAVIS and HECHT)	Social Behaviors, Enforcement, and Tax Compliance Dynamics..	39
PHILLIPS, JOHN (and PINCUS and REGO)	Earnings Management: New Evidence Based on Deferred Tax Expense..............................	491
PHILLIPS, JOHN D.	Corporate Tax-Planning Effectiveness: The Role of Compensation-Based Incentives	847
PINCUS, MORTON (and PHILLIPS and REGO)	Earnings Management: New Evidence Based on Deferred Tax Expense..............................	491

PLUMLEE, MARLENE A.	The Effect of Information Complexity on Analysts' Use of That Information	275
REGO, SONJA OLHOFT (and PHILLIPS and PINCUS)	Earnings Management: New Evidence Based on Deferred Tax Expense	491
REES, LYNN (and SWANSON and JUAREZ-VALDES)	The Contribution of Fundamental Analysis after a Currency Devaluation	875
RYAN, STEPHEN G. (and ZAROWIN)	Why Has the Contemporaneous Linear Returns-Earnings Relation Declined?	523
SEDATOLE, KAREN L.	The Effect of Measurement Alternatives on a Nonfinancial Quality Measure's Forward-Looking Properties	555
SMITH, KIMBERLY J. (and BOSCHEN and DURU and GORDON)	Accounting and Stock Price Performance in Dynamic CEO Compensation Arrangements	143
SMITH, REED (and PATTERSON)	Materiality Uncertainty and Earnings Misstatement	819
STOKES, DONALD J. (and FERGUSON and FRANCIS)	The Effects of Firm-Wide and Office-Level Industry Expertise on Audit Pricing	429
SUBRAMANYAM, K. R. (and HEFLIN and ZHANG)	Regulation FD and the Financial Information Environment: Early Evidence	1
SWANSON, EDWARD P. (and REES and JUAREZ-VALDES)	The Contribution of Fundamental Analysis after a Currency Devaluation	875
TAN, HUN-TONG (and NG)	Effects of Authoritative Guidance Availability and Audit Committee Effectiveness on Auditors' Judgments in an Auditor-Client Negotiation Context	801
TAN, HUN-TONG (and TROTMAN)	Reviewer's Responses to Anticipated Stylization Attempts by Preparers of Audit Workpapers	581
THOMAS, WAYNE B. (and PAYNE)	The Implications of Using Stock-Split Adjusted I/B/E/S Data in Empirical Research	1049

TIRAS, SAMUEL L. (and MANRY and WHEATLEY)	The Influence of Interim Auditor Reviews on the Association of Returns with Earnings	251
TOWRY, KRISTY L.	Control in a Teamwork Environment—The Impact of Social Ties on the Effectiveness of Mutual Monitoring Contracts	1069
TROMBLEY, MARK A. (and ALI and HWANG)	Residual-Income-Based Valuation Predicts Future Stock Returns: Evidence on Mispricing vs. Risk Explanations	377
TROTMAN, KEN T. (and TAN)	Reviewer's Responses to Anticipated Stylization Attempts by Preparers of Audit Workpapers	581
TSE, SENYO Y. (and CLEMENT)	Do Investors Respond to Analysts' Forecast Revisions as if Forecast Accuracy Is All That Matters?	227
VERRECCHIA, ROBERT E. (and CORE and GUAY)	Price versus Non-Price Performance Measures in Optimal CEO Compensation Contracts	957
VISVANATHAN, GNANAKUMAR (and KUMAR)	The Information Content of the Deferred Tax Valuation Allowance	471
WAHLEN, JAMES M. (and BAGINSKI)	Residual Income Risk, Intrinsic Values, and Share Prices	327
WALTHER, BEVERLY R. (and BONNER and YOUNG)	Sophistication-Related Differences in Investors' Models of the Relative Accuracy of Analysts' Forecast Revisions	679
WEAVER, CONNIE D. (and HODDER and McANALLY)	The Influence of Tax and Nontax Factors on Banks' Choice of Organizational Form	297
WEBER, JOSEPH (and BEATTY)	The Effects of Debt Contracting on Voluntary Accounting Method Change	119
WELKER, MICHAEL (and BHATTACHARYA and DAOUK)	The World Price of Earnings Opacity	641
WHEATLEY, CLARK M. (and MANRY and TIRAS)	The Influence of Interim Auditor Reviews on the Association of Returns with Earnings	251

WHISENANT, J. SCOTT (and FAIRFIELD and YOHN)	Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing...........	353
YOHN, TERI LOMBARDI (and FAIRFIELD and WHISENANT)	Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing...........	353
YOUNG, SUSAN M. (and BONNER and WALTHER)	Sophistication-Related Differences in Investors' Models of the Relative Accuracy of Analysts' Forecast Revisions	679
ZAROWIN, PAUL A. (and RYAN)	Why Has the Contemporaneous Linear Returns-Earnings Relation Declined?	523
ZHANG, GUOCHANG (and CHEN)	Heterogeneous Investment Opportunities in Multiple-Segment Firms and the Incremental Value Relevance of Segment Accounting Data........................	397
ZHANG, YUAN (and HEFLIN and SUBRAMANYAM)	Regulation FD and the Financial Information Environment: Early Evidence........................	1

OCTOBER 2003 PLACEMENT ADS

The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The ***Review*** *is published on the first day of January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding.*

401 UNIVERSITY OF NOTRE DAME, Department of Accountancy, invites applications for position openings at the Assistant and Associate Professor levels. All research areas will be considered. Applicants for Assistant Professor positions should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong commitment to scholarly research and excellence in teaching. Salary is very competitive and research support is excellent. Please send *curriculum vitae,* and an example of scholarly work to Professor Tom Schaefer, Chair, Department of Accountancy, Room 102D, Mendoza College of Business Administration, University of Notre Dame, Notre Dame, IN 46556-5646. The University of Notre Dame is an Equal Opportunity, Affirmative Action Employer.

402 NATIONAL UNIVERSITY OF SINGAPORE (NUS), Business School, Department of Finance and Accounting, invites applications for Accounting tenure-track positions at all ranks. Teaching is undergraduate and M.B.A. courses in Financial and Managerial Accounting. Candidates should possess a Ph.D. degree or be expected to obtain one soon. Selection for Associate or Full Professor ranks will be based on an established publication record in reputable journals and proven teaching excellence. Selection for the Assistant Professor rank will be based on potential for publications in reputable journals and for effective teaching. Applications with full resume and three letters of recommendation should be sent to Associate Professor Trevor Wilkins, Head of Department, Department of Finance and Accounting, National University of Singapore, 1 Business Link, Singapore 117592; or Email: fnbhead@nus.edu.sg.

403 THE UNIVERSITY OF SCRANTON, Kania School of Management, AACSB-accredited, is seeking qualified applicants for a tenure-track position beginning August 2004. The university offers courses at the undergraduate and M.B.A. levels. The primary teaching responsibility will be Taxation. Minimum qualifications include an appropriate earned doctorate or ABD progressing toward completion. The University of Scranton in northeastern Pennsylvania is one of 28 Jesuit colleges and universities in the United States committed to providing liberal arts education and strong professional and preprofessional programs in the context of Ignatian educational principles. The University is proud of its mission in the Catholic and Jesuit tradition and spirit, and the successful candidate must be able to support this mission through his or her work in Accounting. Applications will be received until the position is filled. Please send ***curriculum vitae*** to Dr. Michael O. Mensah, Chair, Department of Accounting, University of Scranton, Scranton, PA 18510-4602; Email: gownleyn1@scranton.edu. The University is an Affirmative Action/Equal Opportunity Employer/Educator; women and minorities are encouraged to apply.

404 UNIVERSITY OF MISSOURI–ST. LOUIS, College of Business Administration, invites applications for the position of Assistant or Associate Professor of Accounting, beginning August 2004. Applicants should possess a doctorate or be near completion. There is a strong preference for candidates with interest in teaching Auditing on the undergraduate and graduate levels. The College offers undergraduate and master's degree programs including a B.S. in Accounting, a Master's of Business Administration, and a Master's of Accounting. All programs are AACSB-accredited. The successful candidate should have a strong commitment to excellence in teaching and research. Interested candidates should contact Mary Beth Mohrman, Associate Professor and Coordinator of Accounting, College of Business Administration, University of Missouri–St. Louis, 8001 Natural Bridge Road, St. Louis MO 63121; Email: mohrman@umsl.edu. The University is an Affirmative Action/Equal Opportunity Employer committed to excellence through diversity.

405 KENT STATE UNIVERSITY, Department of Accounting, invites applications for an Assistant and Associate Professor with teaching and research emphasis in Financial or Managerial Accounting beginning in Spring or Fall 2004. A Ph.D. with a major in Accounting or an appropriate related field is required; candidates who are ABD will be considered at the Assistant level. Professional certification such as CPA or CMA is highly desirable. Evidence of the potential for effective undergraduate and graduate teaching, and discovery/integration research is required. Research will be expected in high-quality, academic refereed journals for reappointment, tenure, and promotion; some publication in other journals may be included in the faculty member's portfolio. Participation in teaching doctoral students and serving on and/or supervising their dissertations likely will be part of the faculty member's responsibilities. Candidates at the Associate level must have an established scholarly record. Review of applications will begin September 1, 2003 and continue until the positions are filled. If you are interested in this position, please contact and/or mail a copy of your resume to Dr. Pervaiz Alam, Chair, Faculty Search Committee, Department of Accounting, 576 BSA, Kent State University, Kent, OH 44242; Phone: (330) 672-1121; Fax: (330) 672-2548; Email: palam@bsa3.kent.edu.

406 JAMES MADISON UNIVERSITY invites applications for Director of the AACSB-accredited School of Accounting. Accounting doctoral degree, strong leadership, and interpersonal skills are required. Previous administrative experience is preferred, but not required. Director must promote teaching, scholarship, service, and collegial academic environment. Director must enhance existing school strengths including high-quality students, strong alumni, and employer support. For additional information, contact Dr. Mike Riordan, Accounting Professor, Phone: (540) 568-3209; Email: riordamp@jmu.edu; or Dr. Robert Reid, Dean, College of Business, (540) 568-3254; Email: reidrd@jmu.edu. Application must include *curriculum vitae,* including email and mailing address, contact information for three references, and letter describing ability to meet stated qualifications. Send application to Dr. Phil DuBose, Search Chair, James Madison University, MSC 0205, Harrisonburg, VA 22807; Email: dubosepb@jmu.edu. Application review begins August 25. James Madison University is an Equal Opportunity/Affirmative Action/Equal Access Employer.

407 UNIVERSITY OF NEW HAMPSHIRE, Whittemore School of Business and Economics, invites applications a tenure-track position in Accounting at the Assistant Professor level beginning Fall 2004. Applicants must have a doctoral degree or be near completion. Teaching interests in Financial and Managerial Accounting at both the undergraduate and graduate levels are preferred. The successful candidate will be expected to maintain an active research and publication agenda in peer-reviewed journals. Interested applicants should submit a letter of interest, current *curriculum vitae,* three recommendation letters, recent teaching evaluations, and evidence of scholarly work to Professor Afshad Irani, Whittemore School of Business and Economics, Box A, McConnell Hall, University of New Hampshire, Durham, NH 03824; Telephone: (603) 862-3342. The school is accredited by the AACSB. Application review will begin immediately and continue until the position is filled. UNH is committed to excellence through diversity among its faculty and strongly encourages women and minorities to apply.

408 SINGAPORE MANAGEMENT UNIVERSITY, School of Accountancy, invites applications for Assistant Professor, Associate Professor, or Professor. Faculty interested in visiting appointments are also encouraged to send their *curriculum vitae.* Preferred commencement is July 2004, but negotiable. Preference will be given to candidates who can teach Financial Accounting, Management Accounting, Accounting Theory, and Auditing. All streams of accounting research may be accommodated. CPA is preferred. SMU offers internationally competitive remuneration, research support, reduced teaching loads for research-active faculty, funding to attend international conferences, and relocation assistance and benefits to help international faculty take up their appointments in Singapore. More information about SMU and application procedures are outlined at http://www.smu.edu.sg. Please forward *curriculum vitae* to accountingcv@smu.edu.sg or send to Office of Faculty Administration, Oei Tiong Ham Building, 469 Bukit Timah Road, Singapore 259756.

409 ILLINOIS STATE UNIVERSITY invites applications for tenure-track, Assistant and/or Associate Professor positions in the Department of Accounting, preferably to begin August 16, 2004. Positions are contingent on budgetary approval and appropriate funding. The Department of Accounting at Illinois State University houses both Accounting and Business Information Systems faculty of the College of Business. Ph.D. or D.B.A. with a major in Accounting, completed or advance stages, is required, as well as a strong commitment to quality teaching, service, and research. Initial review of applications will begin September 1, 2003 and continue until position is filled. To assure full consideration, please send a letter of application and evidence of qualifications by September 1, 2003 to James E. Moon, Chairperson, Department of Accounting, Campus Box 5520, Illinois State University, Normal, IL 61790-5520; Email: jemoon@ilstu.edu. Illinois State is an Equal Opportunity, Affirmative Action University encouraging diversity.

410 LEHIGH UNIVERSITY seeks a well-trained scholar who also has a strong commitment to teaching to join its Accounting faculty in Fall 2004, and publish in major refereed journals. The opening is for a Senior Associate or Full Professor with qualifications sufficient to hold the Mercy Professorship. A college with national aspirations, a very successful M.S. in Accounting program, and a dynamic Accounting Department make this a great time to consider Lehigh. Our greatest need is in Financial Accounting. Applications will be accepted until the position is filled. If interested, send your *curriculum vitae,* names of references, and examples of scholarly work to Dr. Kenneth P. Sinclair, Chair, Department of Accounting, Lehigh University, 621 Taylor Street, Bethlehem, PA 18015-3117; Email: kps1@lehigh.edu. Lehigh University is an Equal Employment/Affirmative Action Employer. Qualified women and minority candidates should apply.

411 SOUTHERN ILLINOIS UNIVERSITY EDWARDSVILLE invites applications for a tenure-track position beginning in August 2004. A Ph.D. in Accounting is required, but ABDs nearing completion will be considered. Candidates must be qualified to teach Financial Accounting and must be committed to excellence in teaching and scholarly research. The Department of Accounting offers AACSB-accredited programs at both the undergraduate and master's levels. SIUE is located on a suburban campus of 2,700 acres approximately 25 minutes from downtown St. Louis, MO. Send *curriculum vitae* to Michael L. Costigan, Chair, Department of Accounting, Campus Box 1104, Southern Illinois University Edwardsville, Edwardsville, IL 62026-1104; Email: mcostig@siue.edu. SIUE is an Equal Opportunity/Affirmative Action Employer, and applications from women and minorities are especially encouraged. SIUE is a state university—benefits under state-sponsored plans will not be available to holders of F1 or J1 visas.

412 UNIVERSITY OF MARYLAND, University College Graduate School, seeks adjunct faculty for online teaching assignments in Accounting. No relocation needed; teaching can be done from anywhere in the world. No prior knowledge needed; we train you for online teaching. These positions begin Fall 2003, Spring 2004, or Fall 2004. Area of specialization is open. Candidates must possess a doctorate in Accounting or in Business with Accounting emphasis, and an excellent teaching record. Experience teaching at the Master's level and/or teaching adult learners is strongly preferred. Applicants should send a *curriculum vitae* and three references as an email attachment to blubich@umuc.edu, or mail to Professor Bruce Lubich, Program Director in Accounting, Graduate School, University of Maryland University College, 3501 University Blvd. East, Adelphi, MD 20783. The University of Maryland is an Equal Opportunity/Affirmative Action Employer.

413 STATE UNIVERSITY OF NEW YORK COLLEGE AT GENESEO, Jones School of Business, extends search to fill a tenure-track vacancy for an Assistant Professor of Accounting/Information Systems. The School seeks to fill this position as of September 2004. Teaching responsibility is in the areas of Managerial Accounting and/or Information Systems, with ability to teach both preferred. Faculty are expected to engage in scholarly activity. Requires Ph.D. (or ABD) in Accounting and a demonstrated commitment to effective teaching. Send letter of application, *curriculum vitae,* evidence of teaching effectiveness, and three recent letters of recommendation to Mary Ellen Zuckerman, Dean, Jones School of Business, South Hall 113, State University of New York College at Geneseo, 1 College Circle, Geneseo, NY 14454; Email: Waller@geneseo.edu. Review of applications begins September 1, 2003. SUNY Geneseo is an Affirmative Action/Equal Opportunity Employer committed to recruiting, supporting, and fostering a diverse community of outstanding faculty, staff, and students.

414 INTERNATIONAL UNIVERSITY OF JAPAN, Graduate School of International Management, considered a top-ranked business school in Japan, will have opening(s) for a faculty position in Accounting starting in April 2004 or later, and is seeking qualified candidates at all ranks. The candidates are expected to hold or in the process of finishing a Ph.D. in Accounting and must have demonstrated record of high-quality teaching and research. Preference will be given to those with interest in Japanese and Asian businesses. Submit a detailed *curriculum vitae* in (Word or PDF format), including the names of three referees, by email to Associate Dean Akitoshi Ito: aito@iuj.ac.jp. International University of Japan, Graduate School of International Management, Yamato-machi, Minamiuonuma-gun, Niigata 949-7277, JAPAN; Phone: +81-257-79-1406; Fax: +81-257-79-1187; Website: http://www.iuj.ac.jp.

415 SAN JOSE STATE UNIVERSITY invites applications for Fall 2004 for a tenure-track Accounting position as Associate or Assistant Professor, or temporary appointment as Visiting Professor as Lecturer with an emphasis in Taxation. The candidate must have demonstrated versatility in subject matter competencies and pedagogy. The teaching assignment is at both graduate and undergraduate levels including traditional and accelerated formats. The position is subject to final approval by the University. The minimum credential requirement is Ph.D. for tenure-track appointments, and an M.B.A. for temporary appointments. Strong preference will be given to candidates who hold a Ph.D. from an AACSB-accredited university. The application deadline is December 31, 2003 or until the position is filled. Send application letter, resume, and three letters of reference to Dr. Joseph Mori, College of Business, San Jose State University, One Washington Square, BT850, San Jose, CA 95192-0066; Telephone (403) 924-3460; Email: mori_j@cob.sjsu.edu. San Jose State University is an Equal Opportunity/Affirmative Action Employer.

416 UNIVERSITY OF WESTERN ONTARIO, Richard Ivey School of Business, seeks candidates for a probationary or tenured appointment in the area of Accounting. The position is available to begin in July 2003. The successful candidate will demonstrate a strong commitment to the practice of management in both research and teaching. Contact the Associate Dean, Faculty Development by telephone: (519) 661-3269 or by email: chiggins@ivey.uwo.ca. This position is subject to budget approval. All qualified candidates are encouraged to apply. However, Canadian citizens and permanent residents will be given priority. The University of Western Ontario is committed to employment equity and welcomes applications from all qualified women and men, including visible minorities, aboriginal people, and persons with disabilities.

417 CONCORDIA UNIVERSITY, the John Molson School of Business, located in beautiful Montreal, Quebec, Canada, invites applications for tenure-track positions in Accountancy at the rank of Assistant, Associate, or Full Professor, effective June 1, 2004. A doctoral degree (completed or near completion) in Accountancy is required. Salaries are competitive and commensurate with qualifications and experience. These positions are subject to budgetary approval. The Department of Accountancy is the only department in Canada accredited by the AACSB-The International Association for Management Education. The department has 26 full-time faculty members, is one of the largest in the country, and has an established record of excellence in professional Accounting education and academic research. Please send your complete application package by December 1, 2003 to Dr. Dominic Peltier-Rivest, Chairman, Department of Accountancy, John Molson School of Business, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8 Canada; Email: drivest@jmsb.concordia.ca; Telephone: (514) 848-2424, ext. 2778. Concordia University has a policy of employment equity, including affirmative action for women faculty. In accordance with Canadian immigration requirements, priority will be given to Canadian citizens and permanent residents of Canada.

418 TEXAS CHRISTIAN UNIVERSITY, Department of Accounting, invites applications for a non-tenured Professor of Professional Practice in Accounting. The position is for a nine-month contract beginning Spring 2004, with eligibility for a long-term renewable contract. Candidates will teach 12 hours of undergraduate and M.B.A. or M.Ac. Accounting courses per semester and have some service responsibilities. Candidates must possess an earned doctorate in Accounting (or a closely related field), and have 3+ years experience in classroom teaching. Excellent teaching at all levels is essential. Preference will be given to candidates possessing subject matter breadth. There are no expectations of research productivity for reappointment. Salary range is competitive and commensurate with qualifications. Send applications to Dr. Robert L. Vigeland, Neeley School of Business, TCU Box 298530, Texas Christian University, Fort Worth, TX 76129. Fort Worth is an attractive and affordable place in which to live. Please visit http://www.tcu.edu for further information on TCU and http://www.neeley.tcu.edu for information on the Neeley School. TCU is an EEO/AA Employer.

419 LOYOLA UNIVERSITY NEW ORLEANS, College of Business, seeks an Assistant Professor of Accounting for a full-time, tenure-track position beginning Fall 2004. Responsibilities: teaching Accounting courses at undergraduate and graduate levels; advising students; college, university, and community service; and engaging in research. Highest consideration given to candidates in Accounting Information Systems and/or Taxation. Minimum qualifications candidacy (ABD) for the Ph.D. in Accounting; holding one or more professional certifications is a plus. Candidates should have an established record of quality teaching and ability to show dedication to the generation of quality research presentations and publications. Please submit a letter of application and resume to Chair of Search Committee, Accounting Position, Loyola University New Orleans, Box 15, New Orleans, LA 70118. Applications considered until position is filled, however, applicants are encouraged to submit applications by September 15, 2003. Loyola is an Equal Opportunity, Affirmative Action Employer.

420 LOYOLA UNIVERSITY NEW ORLEANS, College of Business, seeks an Area Chairperson of Accounting beginning Fall 2004. This is a tenure-track position at the rank of Associate or Full Professor. Responsibilities: Area Chairperson responsibilities; teaching Accounting courses at undergraduate and graduate levels; advising students; college, university, and community service; and engaging in research. Also will lead the Accounting area in its efforts to achieve initial accounting accreditation; review year for accounting accreditation is 2006–07. Minimum qualifications: administrative experience and an established record of quality teaching and research. Send application letter and resume to Chair of Search Committee, Accounting Position, Loyola University New Orleans CBA, 6363 St. Charles Ave., Box 15, New Orleans, LA 70118. Applications will be considered until the position is filled. Applicants are encouraged to submit applications by September 15, 2003 for full consideration. Loyola is an Equal Opportunity, Affirmative Action Employer.

421 UNIVERSITY OF COLORADO AT DENVER seeks applicants for a tenure-track faculty position in Accounting. Minimum qualifications include a doctoral degree in Accounting and commitment to excellence in research and teaching. Candidates for advanced rank must demonstrate significant accomplishments in research, teaching, and service. Tax specialization is preferred, though all areas of Accounting will be considered. Proficiency with Information Technology in teaching and/ or research applications is especially desirable. Position is dependent on funding. Review of applications begins immediately, and will continue until the position is filled. Please send a letter of application, current *curriculum vitae,* relevant publications or working papers, and teaching credentials to Professor Bruce R. Neumann, Director of Accounting Programs, University of Colorado at Denver, Campus Box 165, PO Box 173364, Denver, CO 80217; Fax: (303) 556-5899; Email: bruce.neumann@cudenver.edu. The University of Colorado at Denver is an Equal Opportunity Employer and Educator committed to excellence through inclusiveness. The Colorado Open Records Act applies to all application materials.

422 SIENA COLLEGE, Accounting Department, invites applications for a tenure-track position beginning Fall Semester 2004. Applications will be accepted until the position is filled. Rank and salary are dependent upon experience and credentials. Relevant business experience and professional licensing preferred. A Ph.D. or a D.B.A. with a concentration in Accounting is expected. The candidate should be able to demonstrate a track record of teaching excellence and a record of professional development activities. Experience and research interests in Financial or Managerial Accounting, are highly desirable. Please forward applications to Accounting Search Committee, c/o School of Business, Siena College, 515 Loudon Road, Loudonville, NY 12211-1462. Siena College is an independent liberal arts college in the Catholic and Franciscan tradition, located in New York's Capital District, in close proximity to other major higher education and cultural institutions. The college enrolls 2,800 full-time students, 1,200 in the School of Business. The Business School is currently in candidacy for AACSB accreditation. Siena College is an Equal Opportunity Employer and encourages applications from all qualified candidates.

423 CALIFORNIA POLYTECHNIC STATE UNIVERSITY, San Luis Obispo, Accounting Area, Orfalea College of Business, invites applications for full-time tenure-track Assistant/Associate Professor positions in primary areas of Accounting Information Systems, Financial Accounting, Managerial Accounting, and/or Professional Ethics for Accountants. Position begins September 13, 2004. Salary will be commensurate with qualifications and experience. Candidate must have doctorate or equivalent degree in Accounting or closely related field from AACSB or equivalently accredited institution. Applicants must demonstrate a strong teaching and research record and currency or potential in professional field. To apply, visit http://www.calpolyjobs.org and complete the online application. Submit application to Requisition #100109. Please attach letter of interest, current *curriculum vitae,* evidence of teaching excellence, and three letters of recommendation. Letters and student evaluations may be faxed to Mary Beth Armstrong at (805) 756-6347. Cal Poly is strongly committed to achieving excellence through cultural diversity. EEO.

424 STATE UNIVERSITY OF NEW YORK AT BUFFALO expects to have one or more faculty openings for Fall 2004, subject to budgetary approval. Rank and area are open; we are actively seeking senior applicants. Must have demonstrated capacity to undertake scholarly research publishable in leading journals and to provide quality teaching at the undergraduate, M.B.A., and/or Ph.D. level. Doctorate in hand or (at the Assistant level) near completion from a reputable research university is expected. Research record is a requirement for senior positions. Salary and fringe benefits are competitive. Send *curriculum vitae,* two letters of recommendation, copy of dissertation proposal, and/or other research paper(s) to Professor Susan Hamlen, Chair, Department of Accounting and Law, Jacobs Management Center, State University of New York at Buffalo, Buffalo, NY 14260-4000; Fax: (716) 645-3823; Email: hamlen@buffalo.edu. SUNYAB is an Affirmative Action/Equal Opportunity Employer.

425 MCMASTER UNIVERSITY, DeGroote School of Business, a leading business school in Canada (#1 in papers per capita, http://www.bus.ualberta.ca/citationstudy2), is seeking tenure-stream faculty in Accounting at the Assistant Professor level. We are looking to fill one position in July 1, 2004 and possibly one more in July 1, 2005. Applicants should have primary teaching and research interests in Financial/Managerial Accounting or Assurance/Information Systems. Preferred candidates will possess a completed doctorate or be close to completion; exhibit research excellence (or potential) and display teaching excellence (or effectiveness). Please send a *curriculum vitae,* official transcripts, three letters of reference, teaching evaluations (if any), and a working paper to Professor S. M. Khalid Nainar, Chair, Accounting and Financial Management Services Area, DeGroote School of Business, McMaster University, 1280 Main Street West, Hamilton, Ontario, L8S 4M4, CANADA. Competitive salary and benefits; and an attractive teaching load. For further information, please email nainar@mcmaster.ca or visit http://www.degroote.mcmaster.ca.

426 XAVIER UNIVERSITY, a Catholic university in the Jesuit tradition, invites applications for a tenure-track position at the Assistant or Associate level beginning Fall 2004. A Ph.D. in Accounting from an AACSB-accredited program is required. ABDs will be considered. Teaching assignments include Accounting Information Systems as well as Accounting Principles, with a preference for Managerial Accounting. A strong commitment to high-quality teaching and a research agenda that will lead to publication in refereed journals is required. Xavier University's Williams College of Business, which is AACSB-accredited, offers an undergraduate major in Accounting and an M.B.A. Please direct a letter of application and *curriculum vitae* to Edward J. VanDerbeck, Chair, Department of Accountancy, Xavier University, Cincinnati, OH 45207, or email to vanderbeck@xavier.edu. Deadline for all applications is November 1, 2003. Xavier University is an Equal Opportunity/Affirmative Action Employer.

427 PURDUE UNIVERSITY, the Krannert Graduate School of Management, invites applicants for one or more tenure-track appointments in Accounting beginning Fall 2004. Rank is at either the Assistant or untenured Associate Professor level. Candidates for an Assistant Professor position should have (or be near completion of) a doctorate and exhibit the capacity for research, publishable in top-tier accounting journals, and excellence in teaching. Candidates for an Associate Professor position should have an earned doctorate and an established record of excellence in research, teaching, and service on doctoral dissertation committees. We offer a competitive compensation package. Applications will be reviewed beginning September 15, 2003. Applications should include a resume, representative research output, evidence of teaching ability, and you should be prepared to provide at least three letters of reference. Applications should be sent to Professor Mark Bagnoli, Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907. Purdue University is an Affirmative Action, Equal Opportunity Employer. Women, minorities, and individuals with disabilities are encouraged to apply.

428 BAYLOR UNIVERSITY, Department of Accounting and Business Law, invites applications for tenure-track positions at the Assistant and Associate Professor levels for Fall 2004. All research and teaching areas will be considered. A Ph.D. (or ABD) in Accounting is required; professional certification is preferred. Applicants must demonstrate research and teaching excellence. Applications, including *curriculum vitae,* three letters of reference, working paper, or defended dissertation proposal, and evidence of an active Christian faith commitment should be sent to Dr. Charles E. Davis, Department of Accounting and Business Law, Baylor University, Box 98002, Waco, TX 76798-8002; Email: Charles_Davis@baylor.edu. Applications will be reviewed beginning September 15, 2003 and will be accepted until the positions are filled. Baylor University is a Baptist university affiliated with the Baptist General Convention of Texas. As an Affirmative Action/Equal Employment Opportunity Employer, Baylor University encourages minorities, women, veterans, and persons with disabilities to apply.

429 RENSSELAER POLYTECHNIC INSTITUTE, Lally School of Management and Technology, invites applications for a three-year non-tenure-track position at the Assistant or Associate Professor level available in the Fall Semester 2004, but may begin as early as Spring Semester 2004. Candidates should preferably hold a Ph.D. in Accounting and demonstrate evidence of excellent teaching and active research. This position focuses primarily on teaching, however, candidates with research interests in entrepreneurship and innovation are especially encouraged. Teaching will be at the undergraduate and M.B.A. levels. We offer an excellent benefits package including health, dental, life insurance, retirement, tuition, etc. Visit our website at: http://www.rpi.edu/dept/hr. Interested individuals should send a *curriculum vitae* and three letters of recommendation to Ms. Jill Keyes, Accounting Search Coordinator, Lally School of Management and Technology, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, 12180-3590; Email: keyesj@rpi.edu. Review of application material will begin immediately and continue until the position is filled. Rensselaer is an Equal Opportunity/Affirmative Action Employer. Women and minorities are strongly encouraged to apply.

430 RENSSELAER POLYTECHNIC INSTITUTE, Lally School of Management and Technology, invites applications for an Accounting position at the Visiting Faculty or Instructor level beginning Fall Semester 2003. This is a one-year appointment while a search is made for a clinical faculty hire. Candidates should preferably hold a Ph.D. in Accounting and demonstrate evidence of excellent teaching and active research. This position focuses primarily on teaching and appropriate professional experience is a plus. Teaching will be at undergraduate and M.B.A. levels. We offer an excellent benefits package including health, dental, life insurance, retirement, tuition, etc. Visit our website at: http://www.rpi.edu/dept/hr. Interested individuals should send a *curriculum vitae* and three letters of recommendation to Ms. Jill Keyes, Accounting Search Coordinator, Lally School of Management and Technology, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, 12180-3590; Email: keyesj@rpi.edu. Review of application material will begin immediately and continue until the position is filled. Rensselaer is an Equal Opportunity/Affirmative Action Employer. Women and minorities are strongly encouraged to apply.

431 MIAMI UNIVERSITY, Department of Accountancy, invites applicants for a tenure-track position at the Assistant Professor level beginning Fall Semester 2004 (pending funding). Applicants with an interest in Auditing, Financial, Managerial, or Accounting Systems will be considered. The successful applicant will teach undergraduate and graduate courses. Applicants must have a doctorate in Accounting or dissertation in process with expectation of completion of degree by December 31, 2004. Professional certification and experience are desirable. In addition to teaching excellence, faculty members are expected to conduct and publish research. Salary and benefits are competitive. Submit your resume and letters from three references to Dr. Peter Brewer, Department of Accountancy, Miami University, Oxford, Ohio 45056; Email: brewerpc@muohio.edu. Miami University has a national reputation for quality teaching. The undergraduate and graduate programs are AACSB-accredited. Information about the program is available at: http://www.sba.muohio.edu/sba_web/. Miami University is an EEO Employer. Women and minorities are encouraged to apply.

432 VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY, Pamplin College of Business, Department of Accounting and Information Systems, invites applications for an Assistant Professor, tenure-track position starting August 10, 2004. Applicant's primary research and teaching interests should be in Cost/Managerial Accounting. Successful candidate will instruct undergraduate and graduate students in Cost/Managerial Accounting. Qualifications: Ph.D. in Accounting, doctoral requirements completed, demonstrable prospects for conducting scholarly research and excellence in teaching, and business experience in Accounting. Applications will be accepted until October 15, 2003, or until position is filled. Send applications to Robert M. Brown, Department of Accounting and Information Systems, 3007 Pamplin (0101), Virginia Tech, Blacksburg, VA 24061. Virginia Polytechnic Institute and State University is an Equal Opportunity/Affirmative Action Institution.

433 UNIVERSITY OF WASHINGTON, Tacoma, Milgard School of Business (http://www.tacoma.washington.edu/business), announces a faculty position opening for Autumn 2004. Candidates with a record of successful teaching, established research, professional experience, and certification are preferred. A Ph.D. is required. Candidates who can teach in several areas, including Accounting Information Systems, are preferred. To apply, send a letter of application, a *curriculum vitae,* a research stream overview, and a teaching philosophy statement. The position is open until filled. Send application materials to Milgard School of Business Faculty Search, University of Washington, Tacoma, Finance Office Box 358431, 1900 Commerce Street, Tacoma, WA 98402-3100. Contact Person: D. J. Parker, Accounting Search Chair; Email: djparker@u.washington.edu. The University of Washington, an Equal Opportunity/Affirmative Action Employer, is building a culturally diverse faculty and strongly encourages applications from female and minority candidates. All employees must show employment eligibility verification as required by the U.S. Immigration and Naturalization Service.

434 SOUTHWEST TEXAS STATE UNIVERSITY (SWT), Department of Accounting, invites applications for a tenure-track Assistant/Associate position beginning September 1, 2004. This position is subject to availability of funding. Candidates must have a Ph.D./D.B.A. in Accounting from an accredited university or expect completion by August 31, 2004. Candidates must be able to teach Tax and/or Audit. SWT's College of Business Administration offers a B.B.A. in Accounting, Master's of Accountancy (M.Acy.), and a B.B.A./M.Acy. integrated program. All programs are AACSB-accredited. Information about the department is available at: http://www.business.swt.edu/dept/acct/Default.htm. Please send your letter of application, resume, and the names of three references to Dr. Robert Rutledge, Chair of Accounting Faculty Search Committee, Department of Accounting, Derrick Hall 101, Southwest Texas State University, 601 University Drive, San Marcos, TX 78666; Email: rr28@business.swt.edu. Southwest Texas State University is an Equal Opportunity Employer. Women, minorities, and individuals with disabilities are encouraged to apply.

435 UNIVERSITY OF REDLANDS, School of Business, invites applications for a tenure-track faculty position in Accounting for 2004–05. Requirements include an earned doctorate in Accounting (or a Ph.D. in a nonbusiness field with one of the following: one additional earned graduate degree in business; scholarly work in business; or extensive business experience) and clear evidence of research and teaching achievement or potential. All applicants should have a record of achievement in their functional area *and* vibrant interest in cross-disciplinary teaching, adult education, and one or more of the following thematic areas of business: social responsibility, business and the liberal arts and sciences, applied information technology and knowledge management, entrepreneurship and organizational innovation, or leadership and implementation. Applicants should send a cover letter, resume/*curriculum vitae,* copies of transcripts (unofficial copies will suffice), and the names and contact information for three references to Interdisciplinary Faculty Screening Committee, c/o Joanie James, School of Business, Hornby Hall – 101 University of Redlands, 1200 East Colton Avenue, Redlands, CA 92373-0999. Review of applications begins Sept. 1, 2003. Position is open until filled. EOE.

436 UNIVERSITY OF SOUTHERN INDIANA invites applications for tenure-track appointment as Assistant/Associate Professor beginning Fall 2004. Ph.D./D.B.A. desired; ABDs will be considered. Professional certification and relevant experience are desirable. Applicants with interests in Financial/Tax or Financial/Auditing urged to apply. USI is committed to excellence in teaching, scholarship, professional activity, and service to USI and community. Salary and benefits are competitive. The School of Business is AACSB-accredited. Accounting programs also hold separate AACSB accreditation. M.B.A. and M.S.A. programs are supported by Accounting faculty. Application deadline is September 15, 2003; applications accepted until position is filled. Submit letter of application, *curriculum vitae,* and names and addresses of three professional references to Dr. Daniel Wade, Director of Accounting Programs, Department of Accounting and Business Law, University of Southern Indiana, 8600 University Boulevard, Evansville, IN 47712. The University of Southern Indiana is an Affirmative Action/Equal Opportunity Employer.

437 KANSAS STATE UNIVERSITY, Department of Accounting, invites applications for a tenure-track Assistant/Associate position in Taxation beginning August 2004. Position requires teaching undergraduate and graduate courses. Professional experience and certification are desirable. Commitment to teaching excellence and scholarly research leading to refereed publication is required. A proven record of excellence in teaching and scholarship is necessary for appointment at the Associate rank. Candidates should have a doctoral degree in Accounting. ABD near completion will be considered. Review of applications will begin October 15, 2003 and continue until the position is filled. Send a letter of application, *curriculum vitae,* a list of three references, a one-page description of teaching philosophy, and evidence of scholarship to Dr. Dann Fisher, Search Committee Chair, Department of Accounting, Calvin Hall 109, Kansas State University, Manhattan, KS 66506-0502; Email: dfisher@ksu.edu. For more information, please visit http://www.cba.k-state.edu. Kansas State University is an Equal Opportunity Employer and seeks diversity.

438 BABSON COLLEGE, Division of Accounting and Law, invites applications for one full-time tenure-track position in Management Accounting beginning in September 2004. The position will involve teaching at both undergraduate and M.B.A. levels and is open to experienced candidates of all ranks. Applicants should hold a terminal degree and have a strong record in teaching and writing. The strongest candidates will demonstrate practical management or consulting experience, cross-disciplinary interests, and a history of pragmatic, manager-oriented research. Classified as the #1 "hidden gem" business school by *The Wall Street Journal*, Babson's innovative, integrated, and award-winning curricula have garnered high rankings for its undergraduate, graduate, and executive programs. Additional information about the College and the position can be found at http://www.babson.edu. the website provided below. Applicants should send a *curriculum vitae* with cover letter to Search Committee, c/o Dr. Joanne Williams, Chair, Division of Accounting & Law, Babson College, Babson Park, MA 02457-0310. Materials submitted by January 1, 2004 will receive full consideration. Babson College is an Equal Opportunity/Affirmative Action Employer.

439 WESTERN NEW ENGLAND COLLEGE, School of Business, invites applications for a tenure-track Assistant/Associate Professor position in Accounting to begin in September 2004. Candidates are expected to hold a doctorate degree in Accounting or be in the advanced stages of completing their dissertation by the starting date. The successful candidate will demonstrate a strong commitment to both teaching in the undergraduate and graduate programs and to scholarly research. Candidates in all Accounting disciplines will be considered. Salary is competitive and commensurate with credentials. For information about the College, visit our website at http://www.wnec.edu. Apply to Dr. Stanley Kowalski, Jr., Dean, School of Business, Western New England College, 1215 Wilbraham Road, Springfield, MA 01119. All applications should include a *curriculum vitae,* recent teaching evaluations, a recently completed research paper, and the names, addresses, and telephone numbers of three references. Send your inquiry to Professor Vogel at tvogel@wnec.edu. Western New England College is an Equal Opportunity Employer.

440 EAST CAROLINA UNIVERSITY, Department of Accounting, invites applications for an Assistant Professor position in Accounting effective August 15, 2004. Applicants should have/expect to have a Ph.D. in Accounting by December 2004, and appropriate professional certification. Our current needs are Taxation/Managerial. The position is currently funded. ECU is a comprehensive university (22,000 students) with AACSB accreditation. We offer an undergraduate degree and an M.S.A. (concentrations in Taxation and MIS). We are seeking candidates with excellent teaching skills. Appropriate service is also expected. We have reasonable research expectations. Applications will be accepted until the position is filled. Send credentials (cover letter, resume) to Dr. Edwin Doty, Chairperson, Department of Accounting, College of Business, 3208-A Bate Building, ECU, Greenville, NC 27858-4353. Official transcripts are required upon employment. Applicants must comply with the Immigration Reform and Control Act. ECU is an Equal Opportunity/Affirmative Action University and accommodates individuals with disabilities.

441 MISSISSIPPI STATE UNIVERSITY, College of Business and Industry, School of Accountancy, invites applications for Assistant Professor of Accounting, a tenure-track position beginning Spring 2004. Responsibilities will include teaching undergraduate and graduate courses in Accounting Information Systems, Auditing, Tax, and/or Managerial Accounting. Doctorate degree in Accounting is required. ABDs with dissertation near completion may be considered at the instructor level. Requires excellence in teaching, based on prior experience as a graduate Teaching Assistant or as an Instructor or Assistant Professor, ability to do research leading to publication in nationally refereed journals, and an interest and willingness to be involved in professional Accounting activities. An interest in advising and mentoring students is expected. Professional certification and professional Accounting experience are highly desirable. Salary is competitive. Send letter of application, current *curriculum vitae,* and names of three references to Dr. Noel Addy, Associate Professor, School of Accountancy, Drawer EF, Mississippi State, MS 39762. Application will remain open until a suitable candidate is found. MSU is an AA/EOE.

442 UNIVERSITY OF SOUTH FLORIDA, School of Accountancy, invites applicants for the Johnson Endowed Professor of Accounting beginning Fall 2004. Candidates should have an earned doctorate, credentials commensurate with the rank of Associate or full Professor, demonstrated excellence in teaching and research. Preference will be given to candidates with teaching and research experience in Managerial Accounting. To apply or for more complete description of position and minimum qualifications, contact Robert M. Keith; School of Accountancy; College of Business Administration; University of South Florida; 4202 E. Fowler Avenue (BSN3403); Tampa, FL 33620-5500; Phone: (813) 974-6516; Fax: (813) 974-6528; Email: rkeith@coba.usf.edu. Applications close November 7, 2003. Position is subject to final university funding. USF is an Equal Opportunity/Affirmative Action Institution. For disability accommodations, call (813) 974-6389 a minimum of five working days in advance. Under Florida law, applications and meetings are public.

443 UNIVERSITY OF SOUTH FLORIDA, School of Accountancy, invites applicants for tenure-earning Assistant Professor of Accounting beginning Fall 2004. This position will involve teaching in undergraduate and master's programs and involvement with Ph.D. program. Other duties include conducting research leading to publications in refereed journals and service to the academic and professional communities. Ph.D. or near completion is required. Preference will be given to candidates specializing in Financial or Managerial Accounting. To apply or for more complete description of position and minimum qualifications, contact Robert M. Keith; School of Accountancy; College of Business Administration; University of South Florida; 4202 E. Fowler Avenue (BSN3403); Tampa, FL 33620-5500; Phone: (813) 974-6516; Fax: (813) 974-6528; Email: rkeith@coba.usf.edu. Applications close November 7, 2003. Position is subject to final university funding. USF is an Equal Opportunity/Affirmative Action Institution. For disability accommodations, call (813) 974-6389 a minimum of five working days in advance. Under Florida law, applications and meetings are public.

444 APPALACHIAN STATE UNIVERSITY, Department of Accounting, invites applications for tenure-track faculty to teach undergraduate and graduate courses in Auditing, Accounting Info Systems, and/or Financial Accounting. Candidates should have an earned doctorate and have demonstrated potential for quality teaching and research. Preference will be given to candidates with relevant teaching and/or professional business experience and professional certification(s). The start date is August 2004. Applicants must submit a letter of application, current *curriculum vitae,* and a list of three or more references. Review of applications will begin September 15, 2003, and continue until the position is filled. Official university transcripts are required prior to appointment. Send applications to Dr. Randy Edwards, Chairperson, Department of Accounting, ASU PO Box 32013, Appalachian State University, Boone, NC 28608. Appalachian State University is an Equal Opportunity Employer.

445 MICHIGAN TECHNOLOGICAL UNIVERSITY, School of Business and Economics, invites applicants for a full-time tenure-track position in Accounting beginning Fall 2004. A Ph.D. (or ABD near completion) in Accounting from an AACSB-accredited school is required. Applicants in all areas of Accounting are encouraged to apply. Rank and salary will be commensurate with qualifications and experience. Professional certification and experience are a plus. Candidates should have a strong commitment to excellence in teaching, research, and service. Michigan Tech is a nationally recognized technological university and is AACSB-accredited. The University is located on the west side of Michigan's beautiful Upper Peninsula, where publicly accessible land and outdoor recreational opportunities abound. More information may be found at: http://www.sbea.mtu.edu. Interested candidates should forward a letter of application and *curriculum vitae* to Dr. Keith Lantz, Dean, School of Business and Economics, Michigan Technological University, 1400 Townsend Drive, Houghton, Michigan, 49931. MTU is an Equal Opportunity Educational Institution, Equal Opportunity Employer.

446 UNIVERSITY OF NORTH CAROLINA AT GREENSBORO, Bryan School of Business and Economics, Department of Accounting, invites applications for an anticipated position at the Assistant Professor and/or Associate Professor level. Teaching and research interests primarily in Tax Accounting with a Ph.D. or ABD status are required. Review of applications will begin December 31, 2003 and continue until the position is filled. Interested candidates should submit a letter of application, *curriculum vitae,* and the names of three references to Professor C. Edward Arrington, Chair, Search Committee, Bryan School of Business and Economics, UNCG, PO Box 26165, Greensboro, NC 27402-6165. EEO/AA.

447 DREXEL UNIVERSITY, LeBow College of Business, Department of Accounting and Taxation, has an opening for a tenure-track Assistant Professor or Associate Professor position beginning in September 2004. Qualified candidates should have an earned Ph.D. and have an interest in innovative research. Teaching interests should include Managerial Accounting, Financial Accounting, and/or Taxation. Located in Philadelphia's University City area, Drexel's LeBow College of Business is AACSB-accredited at the graduate and undergraduate levels and offers doctorate, M.B.A., Executive M.B.A., and undergraduate degrees. The LeBow College focuses on research and the Department of Accounting and Taxation has a highly productive faculty. Salary, benefits and support are competitive. Send application materials to Seong-Yeon Cho, Recruiting Chair, Department of Accounting and Taxation, LeBow College of Business, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104. Drexel University is an Affirmative Action, Equal Opportunity Employer.

448 DREXEL UNIVERSITY, LeBow College of Business, Department of Accounting and Taxation invites applications for the position of Department Head. Qualified candidates should possess a record of scholarship commensurate with appointment as a tenured Professor of Accounting. Candidates should have strong leadership, interpersonal, and administrative skills and a demonstrated record of excellent teaching and course development. The Department of Accounting and Taxation has 12 full-time faculty members, including one chaired professorship, and is in the process of filling two chaired professorships. Located in Philadelphia's University City area, Drexel's LeBow College of Business is AACSB-accredited at the graduate and undergraduate levels and offers doctorate, M.B.A., Executive M.B.A., and undergraduate degrees. The LeBow College focuses on research and the Department of Accounting and Taxation has a highly productive faculty. Send application materials to Michael Gombola, Department of Accounting and Taxation, LeBow College of Business, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104. Drexel University is an Affirmative Action, Equal Opportunity Employer.

449 COLLEGE OF MOUNT ST. JOSEPH, a Catholic liberal arts college located in suburban Cincinnati, invites applications for a tenure-track position at the Assistant level beginning Fall 2004. Candidates with a Ph.D. in Accounting are preferred; ABDs will be considered. Duties include teaching Accounting Principles, Intermediate Accounting, and other courses in the applicant's area of interest. The College offers undergraduate majors in Accounting, Business and Computer Information Systems, as well as a Master's Degree in Organizational Leadership. A strong commitment to high-quality teaching and a research agenda that will lead to publication in refereed journals is required. Applicants should send a cover letter, *curriculum vitae,* evidence of teaching effectiveness, and should arrange for three letters of recommendation to be sent. Contact: John E. Weiler, Chair, Department of Business Administration, College of Mount St. Joseph, Cincinnati, OH 45233. For more information about the college, see http://www.msj.edu. The College of Mount St. Joseph is an Equal Opportunity Employer.

450 WICHITA STATE UNIVERSITY invites applications for a tenure-eligible position as Assistant or Associate Professor. Required qualifications: a Ph.D. or D.B.A. in Accounting (ABD considered) for Assistant Professor rank; a Ph.D. or D.B.A. in Accounting and a significant research record for Associate Professor rank; teaching and research interests in any area of accounting; commitment to quality teaching and research; commitment to serve and interact with the professional community; experience with diverse populations; and demonstrable commitment to diversity. Desirable qualifications: experience in a professional accounting capacity; professional certification; and teaching and research interests in Accounting Information Systems and/or Financial Accounting and/or Tax Accounting, consistent with the needs of the School of Accountancy. This is a nine-month appointment beginning August 2004. To ensure full consideration, submit applications by October 31, 2003. The position remains open until filled. Salary is competitive (position is contingent upon funding). Send a *curriculum vitae* and names of three professional references to Jeff Bryant, School of Accountancy, Wichita State University, 1845 Fairmount, Wichita, KS 67260-0087; Email: jeffrey.bryant@wichita.edu; Phone: (316) 978-6261. Wichita State University is an Equal Opportunity/Affirmative Action Employer.

451 CHAPMAN UNIVERSITY, Argyros School of Business, invites applications for tenure-track positions in Accounting at the Assistant or Associate Professor rank effective August 2004. Salary will be highly competitive. Applicants should have an earned doctorate by the start of employment and be committed to high-quality scholarly research. Candidates with teaching interest/experience in Financial Accounting and Auditing are preferred. The expected teaching load is two courses per semester. Chapman University is in Orange County, California, consistently listed at the top of "Best Places to Live" in North America. The Argyros School is fully accredited by AACSB and offers bachelor's degrees in Business, Economics, and Accounting as well as the M.B.A. and E.M.B.A. Additional information can be found at http://www.chapman.edu/argyros. Applicants should electronically submit a cover letter, resume, and research sample to Professor Glenn Pfeiffer: asbeaccounting@chapman.edu. Chapman University values diversity and is committed to providing equal opportunities to all individuals.

452 UNIVERSITY OF NEW ORLEANS invites applications for one anticipated Assistant Professor position for Fall 2004. Preference will be given to those with teaching interests in Governmental Accounting, Financial Accounting, and Taxation. Candidates must have the Ph.D. with an Accounting major (or ABD with completion of the Ph.D. by the end of 2004). Salary is negotiable and competitive. The Department offers the undergraduate Accounting degree (over 600 majors) and two Master's degrees (including a Tax Option). The Department has separate AACSB accounting accreditation at both the graduate and undergraduate levels. Candidates should send a resume to Professor Philip J. Harmelink, Chairman, Department of Accounting, College of Business Administration, University of New Orleans, New Orleans, LA 70148; Email: pharmeli@uno.edu. The University of New Orleans is a member of the Louisiana State University System and an Affirmative Action, Equal Opportunity Employer.

453 BUTLER UNIVERSITY invites applications for a tenure-track Accounting position at the Assistant or possibly Associate rank, beginning Fall 2004. Research and teaching interest in Managerial or Financial Accounting are preferred. Applicants should possess a Ph.D., be capable of high-quality teaching, and have an active research program. The ideal candidate will be committed to student learning and excited about teaching in an integrated, experiential curriculum. Butler is a selective, private university set on a wooded 300-acre campus in a residential area just north of Indianapolis. The College of Business is AACSB-accredited and enrolls 700 undergraduate and 325 M.B.A. students. An M.P.Acc. program was also recently approved. Interested candidates should send a letter of application and *curriculum vitae* to Russ Kershaw, College of Business Administration, Butler University, Indianapolis, IN 46208; Phone: (317) 940-9841; Email: rkershaw@butler.edu. Representatives will be attending the AAA Annual Meeting in Hawaii. Butler University is an AA/EEO Employer.

454 SOUTHEASTERN LOUISIANA UNIVERSITY, Department of Accounting, invites applications for a tenure-track Assistant Professor for Fall 2004. The candidate should have a primary teaching/research interest in Cost/Managerial with a secondary interest in Financial. Qualifications: a Ph.D. or D.B.A. with a concentration in Accounting is required. Professional certification is also required (e.g., CPA, CMA, etc.). Salary is competitive. The Department has separate AACSB accounting accreditation that was reaffirmed in 2003. The position will remain open until filled; however, to guarantee consideration all application materials must be received by October 3, 2003. Candidates should send a letter of application, *curriculum vitae,* three letters of reference, and a copy of ALL transcripts (originals required upon employment) to Rick Simpson, Department of Accounting, SLU 10468, Hammond, LA 70402. SLU is an AA/ADA/EEO Employer.

455 EMPORIA STATE UNIVERSITY, School of Business, Department of Accounting and Computer Systems, seeks a dynamic individual to serve as Chair beginning June 6, 2004. This twelve-month position requires an earned doctorate in Accounting (including Taxation) or CIS, college teaching experience, and a research background. Leadership experience and appropriate certifications are highly desirable. The Department offers majors in Accounting and CIS. The Department administers economics courses commonly taken by business students and includes faculty who teach these courses. The department includes fifteen faculty and 350 majors. ESU's School of Business is accredited by AACSB International. Please send letter of application, *curriculum vitae,* unofficial transcripts, and names, addresses, and phone numbers of three professional references to Dr. John Rich, Associate Dean, School of Business, Emporia State University, 1200 Commercial Street, Campus Box 4039, Emporia, KS 66801-5087; Email: richjohn@emporia.edu; Phone: (620) 341-5225. Screening will begin November 3, 2003, and continue until position is filled. For further information, please visit http://www.emporia.edu. An AA/EOE institution, Emporia State University encourages minorities and women to apply.

456 THE UNIVERSITY OF ALABAMA IN HUNTSVILLE invites applications for a tenure-track Assistant/Associate Professor in Accounting with primary teaching interest in Managerial Accounting. Associate level candidates should possess a Ph.D. in an Accounting-related discipline and an established track record of refereed publications. Assistant level candidates should have Ph.D. or have an anticipated completion date around December 2004, and must demonstrate a potential to produce high-quality research publishable in refereed journals. All applicants must demonstrate a commitment to teaching excellence, participation in professional service activities, and a desire to function as a team member. Previous professional experience and professional certification are desirable, but not required. The College of Administrative Sciences at UAH is AACSB-accredited at the undergraduate and graduate levels and offers a B.S.B.A. in Accounting and a Master's of Accountancy. Send letter of application, *curriculum vitae,* and names and addresses of at least three professional references to Dr. Jatinder N. D. Gupta, Chairperson, Accounting and IS; Email: guptaj@uah.edu; Phone: (256) 824-6593. (Position number 03/04-171: subject to budgetary approval; starts January 2004; open until filled.) The University of Alabama in Huntsville is an AA/EEO Institution.

457 UNIVERSITY OF GEORGIA, J.M. Tull School of Accounting, is seeking applications for an anticipated tenure-track position in Accounting at the Assistant Professor level with teaching responsibilities in Taxation. The anticipated date of appointment is Fall 2004. Candidates for this position should have either a Ph.D. in Accounting or be well along in the dissertation stage. Applicants must have a commitment to excellence in teaching and scholarship. Salary and benefits are competitive. Applications received before December 31, 2003 are assured of consideration. Please submit application materials (current *curriculum vitae* and dissertation proposal or example of scholarly work) to Professor Ben Ayers, Chair, Faculty Recruiting Committee, J.M. Tull School of Accounting, Terry College of Business, 255 Brooks Hall, University of Georgia, Athens, GA 30502. The University of Georgia is an Equal Opportunity/Affirmative Action Institution and strongly encourages applications from women and minorities.

458 MISSISSIPPI STATE UNIVERSITY, College of Business and Industry, School of Accountancy, announces an opening for a tenure-track position teaching in the School of Accountancy beginning January 1, 2004. Appointment is at the Associate Professor rank teaching undergraduate and graduate courses in Managerial Accounting, Accounting Information Systems, or Auditing. The individual who secures this position will be expected to serve on the School's Ph.D. Committee and teach Ph.D. classes. Applicants must have a doctorate degree in Accounting. Position requires excellence in teaching, based on prior experience as Assistant or Associate Professor; ability to conduct research leading to publication in top-level journals. Applicants are expected to advise and mentor students at all levels. Professional certification and professional accounting experience are desirable. Send letter of application, current *curriculum vitae,* and names of three references to Dr. Noel Addy, Associate Professor, School of Accountancy, Drawer EF, Mississippi State, MS 39762. Position will remain open until a suitable candidate is found. Mississippi State University is an A/EO Employer.

459 UNIVERSITY OF WISCONSIN–WHITEWATER, Accounting Department, anticipates an opening for a full-time, tenure-track Assistant/Associate Professor position for Fall 2004. The minimum qualifications are an earned terminal degree (Ph.D. or D.B.A.) in Accounting or a related area (ABD considered), the ability to teach Auditing, an established research record, and professional certification (e.g., CPA, CMA). Responsibilities include teaching undergraduate and/or graduate Accounting courses along with research/scholarly activities, advising, and service. Salary is competitive and negotiable. Review of applications will begin November 28, 2003 and continue until the position is filled. All interested persons should apply in writing to Dr. Alka Arora, Department of Accounting, C5023, University of Wisconsin–Whitewater, 800 West Main Street, Whitewater, WI 53190-1790. For more information, please visit http://www.uww.edu/Adminaff/HR/employmt/employ.htm.

460 UNIVERSITY OF ILLINOIS, College of Business, Department of Accountancy, is recruiting faculty for full-time appointments at all ranks. Candidates at advanced ranks must possess a doctoral degree, and candidates for Assistant Professor positions must possess or be near completion of a doctoral degree. Applicants for senior positions should have a strong research record and the ability to participate in doctoral education and curriculum development at the undergraduate and graduate levels. In addition, nontenure-track lecturer and visiting positions are available. Applicants should submit a resume that lists research publications and courses taught. Compensation is competitive and will be commensurate with the appointment level and applicant's record. To ensure full consideration, applications should be received by November 1, 2003. Applications received thereafter will receive full consideration if there are positions still available. Proposed starting date is August 16, 2004. Correspondence and resumes should be sent to Professor Paul Beck, Department of Accountancy, 293 Wohlers Hall, College of Business, University of Illinois, 1206 S. Sixth Street, Champaign, IL 61820. The University of Illinois is an Affirmative Action, Equal Opportunity Employer.

461 UNIVERSITY OF OREGON, Department of Accounting, invites applications for one or more tenure-track positions effective September 2004. All Accounting areas will be considered. Requirements for Assistant rank include completed or nearly completed Ph.D., and evidence of superior teaching competence or potential. Requirements for senior ranks include demonstrated teaching competence and publication record in *The Accounting Review, Journal of Accounting Research, Journal of Accounting and Economics*, and/or *Contemporary Accounting Research.* Salary is negotiable. Application review begins December 1, 2003 for senior ranks and January 16, 2004 for Assistant Professors, continues until filled. Send letter of interest, *curriculum vitae,* selected copies of research papers, statement of teaching interests, evidence of teaching performance, and three letters of reference to Accounting Faculty Search Coordinator, Lundquist College of Business, University of Oregon, Eugene, OR 97403-1208; Fax: (541) 346-3341. The University of Oregon is an AA/EO/ADA institution committed to cultural diversity.

462 BARUCH COLLEGE, the Stan Ross Department of Accountancy, invites applications for anticipated positions in Accounting and Taxation for tenure-track positions at all ranks including full Professor level. Rank and salary are dependent on qualifications. Candidates should have a doctoral degree in Accounting or be close to completion of that degree. For ranks above Assistant Professor, the requirements are a completed Ph.D., record of publication, and demonstrated teaching effectiveness. Please submit *curriculum vitae* and a proposal for a research presentation by February 20, 2004, to Professor Bharat Sarath, Baruch College, 17th Lexington Avenue, Box B12225, New York, NY10010. Baruch College is an AA/EO/IRCA/ADA Employer.

United States Postal Service

Statement of Ownership, Management, and Circulation

1. Publication Title The 'Accounting Review	2. Publication Number 0 0 4 - 0 0 0	3. Filing Date 8/30/03
4. Issue Frequency 4 times a year in January, April, July and October	5. Number of Issues Published Annually 4	6. Annual Subscription Price $175

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4)*

5717 Bessie Drive
Sarasota, Fl. 34233-2399

Contact Person
Tracey Sutherland
Telephone
(941) 921-7747

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

5717 Bessie Drive
Sarasota, Fl. 34233-2399

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*

American Accounting Association
5717 Bessie Drive
Sarasota, Fl. 34233-2399

Editor *(Name and complete mailing address)*

Terry Shevlin - University of Washington
231 Mackenzie Hall/SOBA
Department of Accounting Box 353200, Seattle, Wa. 98195-3200

Managing Editor *(Name and complete mailing address)*

Same as above

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	Complete Mailing Address
American Accounting Association	5717 Bessie Drive Sarasota, Fl. 34233-2399

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☒ None

Full Name	Complete Mailing Address

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*
The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:
☒ Has Not Changed During Preceding 12 Months
☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

PS Form 3526, October 1999 *(See Instructions on Reverse)*

13. Publication Title The Accounting Review		14. Issue Date for Circulation Data Below July 2003	
15.	Extent and Nature of Circulation	Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a.	Total Number of Copies *(Net press run)*	9845	9753
b. Paid and/or Requested Circulation	(1) Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541. *(Include advertiser's proof and exchange copies)*	7168	7564
	(2) Paid In-County Subscriptions Stated on Form 3541 *(Include advertiser's proof and exchange copies)*	–	–
	(3) Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution	–	–
	(4) Other Classes Mailed Through the USPS	–	–
c.	Total Paid and/or Requested Circulation *[Sum of 15b. (1), (2),(3),and (4)]* ▶	7168	7564
d. Free Distribution by Mail *(Samples, complimentary, and other free)*	(1) Outside-County as Stated on Form 3541	1613	904
	(2) In-County as Stated on Form 3541	–	–
	(3) Other Classes Mailed Through the USPS	–	–
e.	Free Distribution Outside the Mail *(Carriers or other means)*	–	–
f.	Total Free Distribution *(Sum of 15d. and 15e.)* ▶	1613	904
g.	Total Distribution *(Sum of 15c. and 15f)* ▶	8781	8468
h.	Copies not Distributed	1064	1285
i.	Total *(Sum of 15g. and h.)* ▶	9845	9753
j.	Percent Paid and/or Requested Circulation *(15c. divided by 15g. times 100)*	82%	89%

16. Publication of Statement of Ownership

☒ Publication required. Will be printed in the October 2003 issue of this publication. ☐ Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner

Tracey E Sutherland, Executive Director

Date 9/26/03

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

Instructions to Publishers

1. Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed form for your records.
2. In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation for whom the trustee is acting. Also include the names and addresses of individuals who are stockholders who own or hold 1 percent or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use blank sheets if more space is required.
3. Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
4. Item 15h., Copies not Distributed, must include (1) newsstand copies originally stated on Form 3541, and returned to the publisher, (2) estimated returns from news agents, and (3), copies for office use, leftovers, spoiled, and all other copies not distributed.
5. If the publication had Periodicals authorization as a general or requester publication, this Statement of Ownership, Management, and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during October, the first issue printed after October.
6. In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
7. Item 17 must be signed.

Failure to file or publish a statement of ownership may lead to suspension of Periodicals authorization.